

University of New Orleans

ScholarWorks@UNO

University of New Orleans Theses and
Dissertations

Dissertations and Theses

Fall 12-20-2018

Singing Their Stories: A Musical Narrative of Teaching and Testing

Desi Richter

University of New Orleans, New Orleans, brichter@uno.edu

Follow this and additional works at: <https://scholarworks.uno.edu/td>



Part of the [Curriculum and Instruction Commons](#), [Curriculum and Social Inquiry Commons](#), [Educational Assessment, Evaluation, and Research Commons](#), and the [Teacher Education and Professional Development Commons](#)

Recommended Citation

Richter, Desi, "Singing Their Stories: A Musical Narrative of Teaching and Testing" (2018). *University of New Orleans Theses and Dissertations*. 2532.

<https://scholarworks.uno.edu/td/2532>

This Dissertation is protected by copyright and/or related rights. It has been brought to you by ScholarWorks@UNO with permission from the rights-holder(s). You are free to use this Dissertation in any way that is permitted by the copyright and related rights legislation that applies to your use. For other uses you need to obtain permission from the rights-holder(s) directly, unless additional rights are indicated by a Creative Commons license in the record and/or on the work itself.

This Dissertation has been accepted for inclusion in University of New Orleans Theses and Dissertations by an authorized administrator of ScholarWorks@UNO. For more information, please contact scholarworks@uno.edu.

Singing Their Stories: A Musical Narrative of Teaching and Testing

A Dissertation

Submitted to the Graduate Faculty of the
University of New Orleans
in partial fulfillment of the
requirements for the degree of

Doctor of Philosophy
in
Curriculum and Instruction

by

Desi Richter

B.M.E. Otterbein University, 1996

December, 2018

Dedication

To my kids, Ana, Moses, Abby and Rudy: You, kids, are the music. Anything I could ever write pales in comparison to the art of your lives. You are my joy.

To tireless teachers everywhere doing the good work. This one is for you.

Acknowledgements

My heart swells with gratitude to so many of you who have buoyed me through the time-consuming process of conceiving, composing, and performing this research.

Dr. Austin, four years ago, you saw me. You welcomed me into a program that has been nothing short of life changing. This dissertation would not be arts-based if you had not kept current of emerging trends in research and planted the seed of conducting alternative research. Because of you, I have been able to create research that is not only relevant and hopefully moving. I have been able to do so in ways that are authentic to me. For the countless hours you have spent reading and responding to my many, many words, *thank you*.

To Abram Himelstein, Dr. Farizo, and Dr. Ward: Thank you for your unwavering support in the form of feedback that has encouraged me to sharpen my thinking and to persevere through the sometimes arduous process of framing and executing this research. You have encouraged, nurtured, and pressed. In short, you are the committee of my dreams. The combined force of your expertise and the disposition with which you have wielded it awes me. Again. Thank you.

To my parents, Fred and Gayle Billy: They say values are caught not taught. Whoever “they” are, they are right. Through your lives, you have taught me to be curious, playful and, perhaps most importantly, to honor and value all work as a noble task. I love you both, and it means the world to me that you are here.

To my colleague and partner in dissertating Dr. Jennie Moctezuma: You blazed a bright path, Dr. M. Every day I am in awe of your brilliance, your heart, and your determination. Thank you for listening to my musings, for sharing your insights. On the days that it was difficult to put my body in a chair and keep going, I channeled you. I came to UNO to get a Ph.D. I got a lifelong friend, too. I’d say that’s a pretty wonderful deal.

To my comrades across the hall in educational administration, Shante Williams, Ebony Blackwell, Tim Sattler: I enrolled in educational administration classes to learn how administrators think, to avoid narrowing my own curriculum in ways that would silo my thoughts. I had no idea I would find educators with whom I would enjoy professional camaraderie and whose community would encourage me and nurture my teaching heart. Thank you.

To my dear friends Bethany Jones, Shannon Blady, and Genevieve Rheams: You have encouraged, listened, held my hand when I needed it. I am beyond grateful for you all.

Thank you to my study participants who shared their time, thoughts, and hearts with me in such generous ways. To you and all the teachers who show up every single day to nurture the hearts and minds of all students, I hold you in the highest esteem.

To Guy Tem, my former teacher and now friend: Special thanks to you for the countless cappuccinos and especially for the use of your incredible studio and recording equipment.

To all who listened to this performance and shared your perspectives, thank you for your time, your input, and your desire to engage around issues that affect the students we are all charged to care for.

TABLE OF CONTENTS

LIST OF TABLES	viii
ABSTRACT	ix
INTRODUCTION	1
A BRIEF HISTORY OF U.S. STANDARDS-BASED EDUCATIONAL REFORM	2
PEDAGOGICAL EFFECTS OF STANDARDS-BASED ACCOUNTABILITY	4
CURRICULAR NARROWING	4
CURRICULAR PACING	4
ATTENTION TO TESTED STANDARDS	5
TEACHING TO THE TEST	5
CURRICULAR DELIVERY	5
TEACHER RESPONSES TO THE ACCOUNTABILITY MOVEMENT	5
TEACHER RESPONSES AS COGNITIVE DISSONANCE	7
RESEARCH PROBLEM	8
STUDY SIGNIFICANCE	9
RESEARCH PURPOSE	15
RESEARCH QUESTIONS	16
LITERATURE REVIEW	17
ORGANIZATION OF THE LITERATURE REVIEW	17
QUANTITATIVE STUDIES	19
QUALITATIVE STUDIES	25
INTRODUCTION	25
QUALITATIVE STUDIES 2003-2007	26
QUALITATIVE STUDIES 2008-PRESENT	34
NEW ORLEANS CONTEXT	44
PHILOSOPHICAL STATEMENT	51
INTRODUCTION	51
CONSTRUCTIVISM	53
POSTMODERNISM	54
AESTHETIC INTERSUBJECTIVE PARADIGM	55
ONTOLOGICAL PLURALISM	55
EPISTEMOLOGICAL CONSIDERATIONS	56
AXIOLOGICAL CONSIDERATIONS	56
POSITION ON THE ART-SCIENCE CONTINUUM	57
THEORETICAL FRAMEWORK	58
COGNITIVE DISSONANCE THEORY	58
CONNECTION TO STUDY	58
SPECIFIC METHODOLOGICAL APPLICATIONS	59
RESEARCH PRACTICES	61
PARTICIPANTS AND RECRUITMENT	61
DATA COLLECTION	62
SEMI-STRUCTURED INTERVIEWS	62
FOCUS GROUP	63
EMAIL FOLLOW-UP	63

DATA ANALYSIS.....	64
NARRATIVE ANALYSIS.....	64
PROBLEM-SOLUTION CODING	64
ARTS-BASED CODING	64
SONGWRITING	65
DATA REPRESENTATION	65
EVALUATION	67
RESEARCHER POSITIONALITY	68
A/R/TOGRAPHER.....	68
INSIDER-OUTSIDER	69
ETHICAL CONSIDERATIONS	69
LIMITATIONS	72
REFERENCES	73
VITA	87

LIST OF TABLES

TABLE 1: EVALUATIVE CRITERIA FOR ARTS-BASED RESEARCH PROJECTS	68
---	----

Abstract

This musical, arts-based educational research describes the lived experiences of four K-12 New Orleans educators who believe that end-of-year standardized tests hinder their ability to teach in ways they believe are best. Using songwriting as a form of data elicitation and narrative restorying, this study documents the lived experiences of teachers who have experienced test-related cognitive dissonance. While curricular narrowing and other test-related practices have been studied in many contexts, the perspectives of New Orleans teachers are barely documented. Thus, this study fills a content gap in the testing literature. Musically restorying the data contributes to the accountability literature in three main ways. First, restorying the data as song renders the findings evocatively — that is, in ways that capture the emotion with which the data was originally imbued. Second, because this study is performative (the results were sung live in the community), the opportunity exists to ignite a local conversation aimed at helping teachers navigate testing/teaching conundrums. Finally, as music is one of the least utilized forms of art-based research, this study fills a methodological gap in the arts-based research repository.

Keywords: arts-based research, arts-based educational research, curricular narrowing, standardized testing, New Orleans school reform, musical inquiry, narrative inquiry

Singing Their Stories:

A Musical Narrative of Testing and Teaching

Introduction

"Now that testing is over, I can teach what I want," recently proclaimed a graduate student enrolled in a local university M.Ed. program. This teacher's statement resonated with twin tones of exultation and exasperation — exultation because, as I later learned in an interview, the teacher finally felt freed from the narrow curriculum she was required to deliver prior to end-of-year, standardized testing. Her exasperation resulted from feeling constrained by the curricular content as well as by the pace with which she was required to deliver it.

This teacher, a certified, highly qualified second grade educator, believed in the professional training she had received at her accredited teacher-training program. She had come to embrace a set of beliefs and practices that she believed served her students well. She had read the works of prominent educational theorists — bell hooks, Paulo Freire, Gloria Ladson-Billings, Dewey — and from those works had developed a philosophical stance that she wished to embody through inquiry-based, collaborative learning practices. This teacher wanted her students to own their education. Yet, she believed the practices she was asked to enact did not mesh with her beliefs that education belongs to students and that collaborative, culturally-relevant instructional approaches, which followed the interests of a class, serve students better than curricula aimed at helping students navigate the end-of-year computer-based testing required by her school.

When this teacher exclaimed, "I can teach what I want," she meant that for a few brief weeks following the administration of mandated end-of-year testing, she could teach in the ways she thought were pedagogically appropriate. This teacher's notions of sound practice, grounded

in theory, informed her beliefs about her teacher identity, her role in the classroom, and the instructional practices she desired to implement; yet she felt constricted in her ability to embody what she viewed as a calling for which she was aptly prepared.

The pedagogical frustrations bemoaned by this teacher are not hers alone. Her sentiments echo those of her peers across the nation. Far from being disconnected from the national conversation about school accountability, her concerns are well documented in the standards-based-reform educational literature, and they are deeply nested in a history of standards-based reform.

A Brief History of U.S. Standards-based Educational Reform

Standards-based reform is rooted in longstanding concerns that, compared to their international counterparts, U.S. students chronically underperform. The issue came to a head in 1957 when Russia gained a leg up in the international space race by launching *Sputnik I*. In the wake of *Sputnik*'s launch, apprehension that U.S. scientists were ill-prepared to compete in the Cold War led to the passage of the National Defense of Education Act of 1958 (NDEA) (National Defense of Education Act of 1958). Via the act, by 1960, the United States had funneled \$1 billion into science curriculum reform ("National Defense Education," 2017). Close on the heels of NDEA, passage of the Elementary and Secondary Education Act of 1965 (ESEA) shifted the legislative focus toward assisting economically-disadvantaged and minority students. ESEA birthed both Title I funds (Elementary and Secondary Education Act of 1965) — meant to help close the achievement gap — as well the widespread use of standards to guide K-12 education. The commitment to standards-based reform was reiterated and its reach extended in 1994 with the passage of the Improving America's Schools Act (IASA). IASA legislation required states to adopt performance standards and assessments but, importantly, it did

something else: IASA effectively tethered meeting those standards to the receipt of Title I funds (Improving America's Schools Act of 1994). The upshot of IASA was the accountability-based era in which the United States currently operates.

The No Child Left behind Act of 2002 (NCLB) further specified and codified a set of accountability measures, most of which are still currently in place. Under NCLB, schools were held accountable for students' adequate yearly progress (AYP) (No Child Left behind Act of 2002). In order to monitor AYP, NCLB required schools to administer standardized tests in English, math, and (eventually) science to students in grades three through eight. Louisiana, the state in which this research was performed, also tests third through eighth grade students in social studies ("Grades 3-8 Assessments," 2017). NCLB also mandated English and math assessments at least once in high school. In Louisiana, high school students are currently assessed using the following tests: End-of-Course exams (EOCs) in English II, Biology, U.S. History, LEAP 2025 in English I & II, Algebra I, Geometry, U.S. History, American College Test (ACT), WorkKeys, Advanced Placement (AP) Tests, College Level Examination Program (CLEP), the Louisiana Educational Assessment Program (LEAP) Connect, and the English Language Proficiency Test (ELPT) ("High School Assessments," 2017).

Schools who did not bring adequate numbers of students to proficiency on these tests faced both fiscal repercussions and possible takeover. Because of the magnitude of the impact inadequate test scores can have on schools, they have come to be called high-stakes tests. Even though the most recent legislations, The Every Student Succeeds Act (ESSA) returns some autonomy to states by allowing them to choose their own standards (Every Student Succeeds Act of 2015), ESSA still links federal funding to standardized test results. Public American education essentially remains, at the time of this writing, accountability-based.

Pedagogical Effects of Standards-Based Accountability

Examining the rhetoric of NCLB reveals passages replete with noble intention. Take, for example, the stated purpose of NCLB, “to ensure that all children have a fair, equal, and significant opportunity to obtain a high-quality education” (No Child Left Behind Act of 2002, Sec 1001, para 1). Or consider some of the specific ways by which NCLB proposed to accomplish that goal, such as by targeting “the educational needs of low-achieving children in our Nation's highest-poverty schools, limited English proficient children, migratory children, children with disabilities, Indian children, neglected or delinquent children, and young children in need of reading assistance” (No Child Left Behind Act of 2002, Sec 1001, para 3).

If the aims of the NCLB are altruistic, the specific ways in which states, districts, and schools implemented the mandate to “improve and strengthen accountability, teaching, and learning” (No Child Left Behind Act of 2002, Sec 1001, para 6) gave rise to a set of testing-related practices that have been documented throughout the country (Au, 2007).

Curricular narrowing. Linking federal funding to performance on high-stakes tests exacerbated the pedagogical phenomenon of curricular narrowing (King & Zucker, 2005; Yeh, 2005). When curricular narrowing occurs, tested subjects (English, math, science) are allotted the lion’s share of class time; conversely, non-tested subjects (social studies, physical education, the arts) are curtailed, sometimes to the point of being eliminated from the curriculum altogether (Amrein & Berliner, 2002; Jerald, 2006; Yeh, 2005).

Curricular pacing. In addition to delivering narrow curricula, teachers who serve in contexts where testing drives instruction have reported feeling pressure to rush curricular delivery (Berliner, 2011; Richter, 2017a; Sondel, 2016). The result of this quick curricular

pacing is that attention is given to breadth over depth of curricular content. Such pacing is deemed necessary in order to cover all the potentially assessed standards.

Attention to tested standards. While all standards may be tested, teachers have described giving more attention to standards most likely to be tested. This attention manifests as spending more time teaching certain standards (Berg, 2006). Similarly, it may involve changing the order of the curriculum so that more heavily tested standards are covered later in the year — closer to testing season so that they are more likely be retained (Blazer, 2011).

Teaching to the test. In addition to the above practices, teachers have also reported spending classroom time teaching students the skills that they need to take the test — that is, teaching to the test (Abrams, Madaus, & Pedulla, 2003; Blazer, 2011; Croco & Costigan, 2007; Jones & Egley, 2004). At times, using simulated test questions, teachers have instructed students about the format of the test, how to think like a test maker, how to answer certain types of questions using test-taking strategies.

Curricular delivery. Two main changes have been documented in the way that teachers have adapted their curricular delivery. First, teachers have adopted lecture-based, direct instruction approaches over more student-centered approaches such as problem-based or inquiry-based learning (Au, 2007). Second, teachers have incorporated test-centric drills into their curricula (i.e. drilling lists of frequently used words, working the types of commonly tested math problems) (Au, 2007).

Teacher Responses to the Accountability Movement

Coexisting alongside the literature documenting post-NCLB curricular changes is a body of research exploring the perspectives of teachers who teach in contexts where the national policy has led them to adopt some or all of the above-outlined practices. Much of this literature

reports on ways in which teachers are dissatisfied with the practices they are either directly or indirectly asked to employ.

Teacher responses to standardized testing phenomena exist along a continuum. This variety is to be expected, given the complex nature of human cognition and the different extents to which accountability metrics interface with regular daily practice. However, in the decade and a half since NCLB became law, teachers have reported varying degrees of discontent with the practices they undertake or the pressure they feel to bring adequate numbers of students to proficiency on end-of-year tests. These findings are explored in greater depth during the literature review, but briefly, teachers have noted feeling stress (Costigan, 2008), anxiety (Sondel, 2016), discouragement (Santoro, 2011) and concerns over the narrow curricula and use of instructional time (Nichols & Berliner, 2008).

In addition to naming the emotions that some teachers have experienced, delving into cognition, the beliefs and mental process related to those emotions, is a useful endeavor because emotions can arise for many reasons. Segall (2012) recognized that “the meaning of standardized testing and its implications for teachers are not pre-determined but, rather, are constructed through teachers' perceptions — thoughts, feelings, beliefs — of them as they interact with the test and its discourses” (p. 287). Emotions are the “whats” of mental phenomena; beliefs and thoughts are the undergirding “why.” Understanding those whys can help educators frame action steps that are appropriate to the cause of the dissatisfactory emotion. For instance, if teachers are frustrated with the pressure to raise test scores, but that frustration arises from a lack of proper training, then one-step toward mitigating that frustration might include engaging in test-preparation-centric, professional development. If the root of the frustration lies elsewhere, then different action steps would be called for.

In the case of test-related frustrations, the teachers' pedagogical discontent seems to stem from their beliefs about what constitutes good teaching. Several studies noted a disconnect between teachers' beliefs about what good teaching is and the actual ways in which they are required to teach (Abrams, Pedulla, & Madaus, 2003; Arroyo-Romano, 2016; Berliner, 2011; Rapp, 2002). Importantly, the teachers in these studies read their frustrations back to the accountability measures (i.e. testing practices) they are required to employ. These teachers believe that the attention they must give to the test directly and negatively impacts their ability to teach in ways they deem best. In other words, teachers in these studies are experiencing a conflict, a "teach to the test or teach best" conundrum.

Teacher responses as cognitive dissonance. In recent years, researchers have begun formulating frameworks that describe teachers' experiences with standardized testing during the age of accountability. Santoro (2011) posited that teachers were experiencing a moral dilemma and that those who left impoverished schools where testing dictated their pedagogy were practicing a type of conscientious objection. Arroyo-Romano (2016) framed teachers' responses to testing-related practices as an ethical dilemma (teachers' ethics dictated certain types of practice; testing dictated another). Both Santoro's and Arroyo-Romano's frameworks are certainly plausible. While they shed light on the forces that could be driving test-related discontent, this study posits that the disparity between teachers' beliefs about pedagogy and the actual test-mediated practices they undertake can be viewed as type of testing-related cognitive dissonance. Taking this view does not negate other previously explored theories, especially when one considers that moral and ethical dilemmas can be viewed as types of cognitive dissonance (Graham, 2007).

When internal beliefs and external actions fail to align, people can experience uncomfortable mental and emotional states. Leon Festinger (1957) called this resulting discomfort cognitive dissonance. If teachers' reported disconnect is viewed as cognitive dissonance, then studying the problem through this lens could yield new psychological insights into their experiences. Moreover, because the theory itself posits avenues of navigating cognitive dissonance, studying teacher discontent as cognitive dissonance could provide real solutions for a subset of struggling teachers. Such resolution is important because, left unchecked, cognitive dissonance can lead to depression, anxiety, and post-traumatic stress disorder (Hull, 2002). For teachers who may be experiencing cognitive dissonance, the mental stakes could be quite high.

Research Problem

In many parts of the nation, the disconnect between how teachers *believe* they should practice and the ways in which they are actually asked to practice is a real and present problem (Costigan, 2008; Croco & Costigan, 2007; Ferguson, Kober & Rentner, 2016; "Primary Sources," 2010; "Primary Sources," 2012). As even this brief outline of the issue intimates and the complete literature review demonstrates, a substantial cadre of teachers, particularly in contexts where curricular responses to accountability measures are most prescriptive, continue to describe their testing-related discontent. Though the problem of practice is well documented, some schools continue to cater to accountability measures in stringent ways. Several authors (Amrein & Berliner, 2002; Berg, 2006; Blazer, 2011; Valli & Buese, 2007) have asserted that accountability-driven curricular effects are more prevalent in schools serving students of color and low-income families.

The extent to which teacher discontent with test-centric practices exists in New Orleans, where this study takes place, has yet to be determined. I have located only a single study that

specifically examined the curriculum of local teachers (Sondel, 2016). Sondel's study confirmed that curricular narrowing and teaching to the test do exist and that teachers in this study felt test-related pressures. However, this study was primarily focused on the observable practices of teachers, not on their opinions of that practice. Local data from my pilot study (Richter, 2017a) demonstrated that some teachers are experiencing the same disconnect documented in the literature, namely that they are encouraged (at times, required) to teach in ways that misalign to their own notions of what constitutes good teaching. The extent to which this local problem exists has yet to be researched. However, considering the extent to which data-driven rhetoric pervades the local discussion and that both Sondel's study and my own pilot demonstrate teacher discontent with testing-related practices, I believe that I have found a research vein that needs to be tapped. Thus, this study is a locally-situated examination of the "teach to the test or teach best" conundrum.

Study Significance

Standardized testing is easily one of the most widely debated topics in American education (Blazer, 2011; Segall, 2003). Proponents of the practice have argued that the testing aligns curricula to standards (Perkins & Wellman, 2008) and holds teachers accountable for their practice (Phelps, 2011). Detractors have pointed to the previously noted pedagogical effects (curricular narrowing, teaching to the test, etc.). In addition, some have argued that the tests tied to the eugenics movement of the early 1900s (Au, 2013) retain a racial bias that is inherently discriminatory (Santelices & Wilson, 2010). The cornerstone test of the Louisiana accountability movement, the Louisiana Educational Assessment Program (LEAP), has certainly not been immune to such criticisms (Decuir, 2012).

Analyzing whether or not standardized tests are viable measurements of student achievement is not the focus of this research. The focus of this research remains squarely on lived experiences of the teacher participants. However, whether or not the gap in achievement that test results purportedly measures is actually closing is decidedly relevant to this study. Closing the gap is the premise upon which accountability measures rest, so whether or not the gap is being closed matters, especially when practices predicated upon closing the gap (such as high-stakes testing) affect the pedagogies of teachers and, in turn, the students they teach. When considering the worth of this study, it is important to consider achievement gap data because one could argue that any teacher frustration could be soothed via ends-means logic. In other words, if teachers are unhappy, then perhaps they need to “take one for the team.” If, on the other hand, the gap is indisputably closing, then perhaps teachers should realign their inner compasses to the necessary practices because, after all, what they are doing is ultimately serving the greater good.

Whether or not the gap is currently being closed is debatable (Amrein & Berliner, 2002; Braun, Chapman, & Vezzu, 2010; Chudowsky, Chudowsky, & Kober, 2009). The most recent data from the National Association of Educational Progress (NAEP) indicate that nationally, in a comparison between the years 2013 and 2015, there have been no significant changes in the Black-White and Hispanic-White achievement gaps in either math or reading (“2015 Math and Reading,” 2017).

In the discussion of whether or not mandated accountability measures have achieved their stated ends — to close the achievement gap — New Orleans is certainly no exception. Locally, accountability proponents argue that reform measures (rooted in school performance scores that rely on end of year testing) are successfully closing the gap (“The Data Story,” 2017). However, upon closer examination, the data from which this conclusion is drawn only report differences

between subgroups in New Orleans and their peers in other areas of the state. In other words, African American students in New Orleans performed better than African American students elsewhere in Louisiana. While this statistic is encouraging, it fails to address whether or not achievement gap between African American (or other minority students) and White students is actually being closed.

According to the Louisiana Board of Elementary and Secondary Education (BESE), that gap remains. In March of 2017, as Louisiana began to implement another round of education reform, ESSA, BESE cited a need for “more vigorous and aggressive policies and strategies to narrow the *persistent and widening achievement gaps* among student subgroups” (“Comments and Questions,” 2017, pg. 2, emphasis mine). The subgroups to which BESE referred are the same as those targeted by NCLB: economically-disadvantaged, minority students, English language learners, students with exceptionalities. In New Orleans, the achievement gap between these students and their counterparts widened between 2013 and 2016 (Dreilinger, 2016) and, last year, the percentage of subgroup K-8 students who obtained mastery level in English, math, and science either decreased or stayed the same (“2015-2017 LEAP Assessment,” 2017). These are the students to whom state superintendent John White referred in 2016 when, referencing state test scores, he stated, “The results lay bare the difficult truth that many students in our state are significantly behind their peers” (Dreilinger, 2016). New Orleans schools where, in 2015, 94% of students were minorities and 84% were economically disadvantaged (“Student Enrollment and Demographics,” 2017), are likely to be the recipients of BESE’s recommended aggressive policies.

Given that New Orleans serves such a preponderance of students targeted by both NCLB and ESSA, given that the success of accountability measures in New Orleans is a tenuous

proposition, and finally, given that the practices tied to passing muster on accountability measures are, at least for some teachers, creating a schism between their notions of good teaching and the teaching they are required to perform, studying the perspectives of these local teachers is a worthy research venture.

As educational stakeholders, the perspectives of teachers should be a thread that is prominently woven throughout the local educational conversational tapestry. Teachers are the ones who deliver curriculum. It is teachers who daily stand face-to-face with the populations the legislation aims to serve. Finally, and importantly, since post-Katrina New Orleans has been touted as ground zero for exemplary school reform (Brinson, Boast, Hassel, & Kingsland, 2012; Chait, 2015), the perspectives of these educators, in particular, should be present in the testing/reform literature. They are not — at least not in any in-depth studies.

In the same year that NCLB became law, Rapp (2002) claimed that the voices of teachers were underrepresented in both the research and legislative decision-making process at large. Rapp described the need for teachers' voices to be engaged with as “pressing” (p. 218). While researchers (Agee, 2004; Augustine, 2016; Bolgatz, 2006; Craig, 2004; Croco & Costigan, 2007; Costigan, 2008; Hinde, 2003, Newberg-Long, 2010; Santoro, 2011; Spohn, 2008; Valli & Buese, 2007) in other parts of the country seemed to answer that call, the actual voices of New Orleans educators are still conspicuously absent from the research. Thus, this research is significant because it fills a rather substantial content gap in the testing literature — voices of educators in a community that has been in flux for over ten years.

The time is ripe for New Orleans teachers' perspectives to be represented not just because they may or may not replicate other study results. As New Orleans sits poised to be the first all-charter city in the United States, perspectives from this unique pool of educators might yield

insights into how reform plays out in an all-charter context. New Orleans teacher voices should also be represented immediately for another reason. In 2012, the Louisiana legislature linked teacher tenure to test score performance. Specifically, any previously tenured teachers, deemed ineffective (as measured by student test scores) can be fired at will. A recent analysis (Strunk, Barrett & Lincove, 2017) studied the effects of the new policy on teacher turnover. Statewide, in the two years since the legislation took force, between 1,500 and 1,700 teachers have exited the profession (an increase of .5% per year). For teachers who practiced in schools where standardized test scores were lowest, the effect was more pronounced. Teacher exodus increased by 27%. The authors proposed that, as the effects of the legislation continue to unfold, improving teacher working conditions may help counter the effects decreased job security Louisiana teachers face. This study sheds light on two aspect of those working conditions: Teachers' perceptions of the curricula they provide and pedagogical practices they employ. Since the authors estimated that Louisiana has already lost 3.0-3.5% of its teacher workforce as a result of the legislation and that tenured or eligible-for-tenure teachers are leaving at higher rates than their less experienced counterparts, *any* research related to teacher satisfaction (like this research) is inevitably timely and significant.

In addition to including the perspectives from a unique (nearly all-charter) context, this study fills a methodological gap in the testing literature. It does so by providing an in-depth, current rendering of teacher perspectives via a narrative approach. In reviewing the past 16 years of testing literature (2002 to present), I have discovered a total of five narrative studies. Two (Hinde, 2003; Craig, 2004) are arguably outdated. While they represent viable pictures of test-driven practices and show that accountability rhetoric greatly informed practices in the early NCLB era, they do not necessarily represent the current state or scope of accountability

legislation. Augustine's (2016) current research is concerned with the ways that high-stakes testing changed classroom dynamics. Certainly this study is relevant to my research as it shows that curricular narrowing and teachers' distaste for it were alive and well in Indiana in 2016. However, neither Augustine nor any other researcher focused on the psychological effects of high-stakes testing narratively in quite the way that I do — through a theoretical lens of cognitive dissonance. Though Alford's (2010) narrative recognized that teachers' frustrations with standardized testing can increase dissonance, test-centric cognitive fallout was not the main thrust of this study. Finally, Ciolino, Kirylo, Mirón, & Frazier's (2014) local study, while certainly useful, provides a broad scope examination of issues that are, for the most part, peripheral to this study. While not wanting to claim too much, it is my firm belief that, at the time of this writing, a New Orleans-specific, narrative study, replete with thick descriptions and first person-participant recountings, is long overdue.

As arts-based research, this study fills twin gaps in the testing literature and the arts-based research repository. Arts-based research is an emerging methodology that involves "the systematic use of the artistic process, the actual making of artistic expressions in all of the different forms of the arts, as a primary way of understanding and examining experience by both researchers and the people that they involve in their studies" (Knowles & Cole, 2008, p. 29). As arts-based research, the art form I used, music, acted as a primary mode of inquiry that was woven through several phases of the research process (Barone & Eisner, 2012; Leavy, 2015; Merriam & Tisdell, 2016). Music acted as a form of data elicitation, analysis, and representation.

No study to date explores testing-related issues using arts-based research as a methodology. Though a single researcher (Newberg-Long, 2010) wrote a set of poetic narratives to accompany her phenomenological findings, her study utilized a different art form. Newberg-

Long self-identified her study as qualitative phenomenology; nowhere did she indicate that arts-based epistemologies informed her methodological decisions. My study, on the other hand, fully embraces arts-based philosophical stances (constructivist, postmodern, aesthetic intersubjective). Thus, it stands as a unique contribution to the field. Finally, even amongst arts-based researchers, musical inquiry and sonic interpretation schema (Leavy, 2015) are less-frequently utilized methods. To date, I have located a handful of studies using music (Bakan, 2014; Carson, 2017; Jenoure, 2002; Viega, 2013). None of them are even remotely related to my research topic. None use songwriting as a way to restory participants' experiences. In this way, my study is not just contributing to arts-based educational research. It is breaking new ground.

When considering this study's significance, one final avenue of thought must be explored. While all published research is in some sense performative, the results of this research were literally performed (sung) before a live audience of New Orleans educational stakeholders. Typically, if researchers are able, they publish their findings in peer-reviewed journals; hopefully, their works are read by other academics and people with an interest in the topic at hand. While I certainly agree with the merits of this practice and plan to publish the results of this study in just such a journal, I also wanted to invite real-time dialogue with those who may not regularly consume research. Local educators and community members heard this research performed and engaged in a conversation after the performance. The results of that discussion were generative and could lead to continued dialogue, community action, and further research.

Research Purpose

This study aims, via the production of musical narratives, to illuminate the perspectives of local (New Orleans) K-12 teachers who practice in contexts where standardized testing impacts their practice. Specifically, the purpose of this study is to narrate the experiences of

these teachers and illuminate the beliefs teachers may have that run counter to the practices they are asked to enact. In this way, as Guest, Namey and Mitchell (2012) explained, this study serves to further the understanding of social phenomena and diagnose psychological conditions (Guest, Namey & Mitchell, 2012). Therefore, as narrative, this study aims to deepen understanding of this socially-situated, educational and psychological issue.

Because this study is arts-based research, it also aims, via emotionally-evocative renderings, to restory data in such a way that those who experience this research can do so vividly and viscerally. Inherent in this particular research agenda is a desire to, “connect audiences on a deeper, more emotional level and . . . evoke compassion, empathy, and sympathy, as well as understanding” (Leavy, 2015, p. 23). While evidence exists (Sondel, 2016; Richter, 2017a) that teachers in New Orleans do experience testing-related cognitive dissonance, regardless of whether or not findings replicate this previous research, I believe that presenting teachers’ voices is worthwhile. Given that as recent as 2016, 94 % of teachers surveyed (n= 10,000) believed that their voices were of no consequence to policy makers (“Primary Sources,” 2012) any research that centers teacher perspectives in evocative ways will likely promote some level of dialogue and, hopefully, viable action.

Research Questions

In order to fulfill the aims of this study, I crafted two research questions. First, I asked, “What are the lived experiences of K-12 New Orleans teachers who believe that standardized testing negatively impacts their teaching?” Since this study seeks to describe the beliefs and attitudes of these teachers, lived experiences encompassed both the external events of their pedagogical experiences as well as teachers’ cognitive and emotional reactions to those events.

I also sought to answer, “If local (New Orleans) K-12 instructors who believe standardized testing negatively affects their practice are experiencing cognitive dissonance, how might they be navigating that dissonance?” These two questions functioned in tandem in order to serve the research agenda. The first question was devised to more broadly capture the experiences of teachers. The second put a finer point on the issue by narrowing the research focus to a specific phenomenon under study: Testing-related cognitive dissonance. I used both questions to frame interview and observation protocols as well as aid data analysis.

Literature Review

Reviewing the standardized testing literature is a bit like drinking from a fire hydrant that has been gushing at full bore — for over twenty years. Even prior to NCLB, researchers documented the effects of standardized testing on teaching practice (Anagnostopolous, 2003a; Anagnostopolous, 2003b; Hillocks, 2002; Luna & Turner, 2001; Taylor, Sheppard, Kinner & Rosenthal, 2001; Smagorinsky, Lakly, & Johnson, 2002; Smith, 2003). After NCLB incentivized testing, research that documented the effects of NCLB on both curriculum and teachers’ perspectives of those effects began to emerge. These two facets of the research, curricular effects and teacher perceptions of those effects, are the main focus of this literature review. While the Seconcernantly worthy of study, they are beyond the scope of this research.

Organization of the Literature Review

I opted to organize this literature review both methodologically and chronologically. I chose this approach for two reasons. First, the quantitative studies pertinent to this research generally have larger samples than their qualitative counterparts. These studies are relevant because they provide a 10,000-foot national view of NCLB implications for practice and teachers responses to those changes. Taken together, these quantitative studies demonstrate the breadth of

the research problem. In contrast, the qualitative studies plumb the depths, the context-specific particulars of pedagogical changes and teachers' perceptions of those changes. Considered side by side, the two methodologies offer complementary views of the problem at hand.

The decision to review studies chronologically is born of two observations. First, since NCLB was implemented, almost annually, new studies have demonstrated similar findings. Reviewing these studies in the order in which they were published allows me to narrate the "plot" of the NCLB-era standardized testing research as it unfolded year by year. The result is a near unbroken line of research documenting how accountability has come to permeate the fabric of the American educational conversation.

A second reason that I chose to write this review chronologically is that I identified a trend in the qualitative research that is mostly clearly seen when the studies are linearly reviewed. The qualitative research occurred in two waves. The first wave (2002 - 2007) documented the curricular effects of NCLB legislation across different contexts. The second (2008 - 2017) focused on understanding how teachers reacted to the effects on their practices. First-wave qualitative researchers seemed intent on describing and documenting the related phenomena of test-related curricular changes and teachers' responses to those changes. Second-wave researchers began to document the beliefs and attitudes of teachers and, in some cases, apply theoretical constructs to teachers' cognitive and emotional processes. This trend in the qualitative data is just that: A general trend. Some early studies address teacher dissatisfaction as an internal/external disconnect (Abrams, Madaus, & Pedulla, 2003). Some later studies re-document earlier findings. The focus, however, shifted away from showing that NCLB had affected practice. Instead, researchers seemed more interested in studying the teachers themselves. Certainly, this is the tack that my research takes, albeit in a way not previously not

applied to the topic (i.e. the use of arts-based research). At the time of this writing, we may be on the cusp of a third wave of research: Studies that document how ESSA affects pedagogies. Thus, my research is potentially situated at the intersection of the “old” (pedagogies enacted under NCLB) and “new” (pedagogies arising under ESSA). Only time will tell how this particular narrative plays out. For now, in order to better understand where we are in the present moment, it is time to turn attention to the past.

Quantitative Studies

Several quantitative studies conducted post-NCLB passage documented both the effects of increased accountability and teachers’ responses to those effects. In 2002, Rapp (2002) published a survey of 191 board-certified teachers in Ohio. Eighty-eight percent of these teachers stated that standardized tests decreased their autonomy in the classroom. In addition, Rapp reported that 91% of the teachers believed that standardized tests fail to support developmentally-appropriate practices. Rapp synthesized the study results with two prior surveys upon which is based (Jones, Jones, Hardin, Yarbrough & Davis, 1999; “Public Agenda Reality Check,” 2000) and concluded that “Each of these studies reinforces the idea that educators believe that classroom instruction is becoming synonymous with test preparation and ultimately leading to losses of autonomy, insight, creativity, and love of learning — for both students and teachers” (p. 216). This study, published eleven months after NCLB became law, is key not because its findings are particularly surprising; in fact, they are echoed nearly yearly by other researchers nationwide. This study is also not unique in that Rapp called for more research that “captures the perspectives of teachers” (p. 218). This study is notable because Rapp specifically called for research that can be used to “engage, provoke, and challenge the public to reclaim its legitimate role in policy formation” (p.218). It is to this second portion of Rapp’s call that my

proposed research, a publicly-performed scholarly work, hearkens, nearly 17 years after it was issued.

One year later, Abrams, Madaus, and Pedulla (2003) reported the results of a large-scale survey of teachers across 30 states (n=4,195). For purposes of the study, the authors delineated states as high-stakes if they attached significant consequences to standardized tests for districts/teachers/schools. After reviewing the stakes attached to various tests, they designated 19 states, including Louisiana, as high-stakes states. In high-stake states, 41% of teachers reported feeling pressure to raise test scores and prepare students for state tests. Slightly more, 43%, reported they taught narrowed curriculum, and over half, 51%, reported teaching to the test. Most disconcerting to these authors was the result that a high percentage of respondents in *all* states (28 in total) indicated that testing programs led teachers to teach in ways that contradict their notions of good educational practice. The percentages in high-stakes states (76%) were only slightly higher than in states deemed low-stakes (73%). The significance of this study lies in both its timing (one year after NCLB was implemented) and its numbers (over 4,000). Even so soon after NCLB was passed, evidence existed that it was affecting practice across the country (even in states without high-stakes attached to tests) in ways that teachers found troubling.

Following on the heels of Abrams, Madaus and Pedulla's work, Jones and Egley (2004) surveyed 708 elementary teachers in Florida. (The study included a single open-ended free-response item, but I include the findings here because, while the researchers did code this open response for themes, this study was primarily quantitative). These researchers believed that teachers' issues with standardized testing were not new, but they wanted to revisit teacher perspectives to see whether or not they had changed two years after NCLB became law. In addition to corroborating curricular narrowing and teaching to the test of previous studies, Jones

and Egley's surveyed teachers reported that testing stifled their creativity and encouraged them to teach curricula that were broad and shallow. Finally, the teachers surveyed indicated that they were not opposed to standards; their issue was the one-size-fits all way in which those standards were assessed.

Jones and Egley's (2004) study yielded findings worth examining and that align with similar studies in the testing literature. However, the authors did not disaggregate the findings according to the demographics of teachers surveyed. Teachers surveyed were categorized according to district type (rural, suburban, and urban). Disaggregating the data by district type would have proved useful for analyzing whether teacher results varied by teaching context. By not disaggregating the data, the ability to correlate teacher perspective by type of district was missed.

The same year as Jones and Egley published, Reese, Gordon, and Price (2004) reported similar results from a statewide survey (n=918) conducted amongst elementary, middle, and high school teachers in Texas. The percentage of teachers overall who emphasized tested material in their instruction (a hallmark of curricular narrowing) was high (70%), but in the elementary grades, where NCLB mandated yearly testing in grades 3-5, the percentages were even higher (90%).

In addition to the independent research above, the Center on Education Policy (CEP) (McMurray, 2007), a nonprofit, nonpartisan, research group, has conducted yearly, national research on the effects of NCLB. In 2007, the CEP surveyed 349 urban, suburban, and rural school districts across all 50 states. Five years after NCLB became law, 62% of the elementary school districts reported an increase in instructional time for English and math. A similar but smaller effect was seen in high school districts, with 20% of districts reporting increased

instructional time in English and math. Concurrently, 44% of elementary districts stated that one or more non-tested subjects had been cut from the curriculum since the 2001-02 school year. This study clearly demonstrated that, at least eight years ago, NCLB had significantly affected the way many districts allotted instructional time, and these changes created a curricular hierarchy. The phenomenon of curricular narrowing was, by this point, a national one. Moreover, the hierarchy extended beyond instructional time given to each subject and into the content covered within tested subjects. Eighty-four percent of elementary districts polled aligned their reading content to the test. The percentages of content-test alignment were similarly high in middle and high school English (79% and 76% respectively). A large preponderance of districts took a similar tack with math; eighty-one% of elementary districts and 71% of high school districts reported weighting post-NCLB curriculum toward tested content.

While the CEP five-year study (McMurrer, 2007) intimated that NCLB wrought national curricular changes in urban, suburban, and rural districts, it did not address how curricular restructuring played out in specific schools and classrooms; nor were teacher responses to those changes addressed. The study did indicate, however, that districts with more struggling schools were more subject to curricular narrowing.

More recently, the Bill and Melinda Gates Foundation published two of the largest (n=40,000 and n=10,000) nationwide surveys (“Primary Sources,” 2010; “Primary Sources,” 2012). Both surveys polled teachers regarding their perspectives on current issues in American education. Most pertinent to this study is the 2010 finding that 90% of teachers valued high-quality curriculum as a mode of ensuring academic success (“Primary Sources,” 2010). The lack of defining terms (i.e. what do teachers consider to be high-quality curriculum) makes it difficult to draw up a set of teacher-preferred curricular action steps based upon this study alone.

However, in 2012, of the 40,000 teachers surveyed, 92% (36,800) indicated that formative, ongoing assessments — the kinds teachers create in their own classrooms — were either essential or very important measures of students’ academic achievement. In contrast, only 27% of teachers believed that data from state-mandated, standardized tests were essential or very important measures of achievement. From these results, it is logical to infer that many, many teachers value the formative, in-class assessments over state-mandated, standardized ones for which they must prepare students. Such a values discrepancy is fertile soil from which teacher frustration and cognitive dissonance could spring.

The 2012 Primary Sources survey (“Primary Sources,” 2012) further teased apart the values disconnect between teachers and policy makers. Similar to the 2009 results, a majority of teachers (92%) indicated that they value in-class, formative assessments (assessment more likely to be teacher-designed) over state-mandated, standardized tests. Moreover, only 26% of these teachers believed that the state tests are accurate reflections of student achievement. Further, the study indicated that a majority of teachers (89%) believed that curriculum that “goes beyond what is tested on standardized tests” (p. 10), would have either a very strong or strong impact on student achievement. As with the previous survey, the study questions failed to designate exactly form such curriculum would take, but it appears that for nearly 8,900 teachers that curriculum would be broader than the curriculum inspired by standardized tests. Considered together, these two studies implied that large swaths of American teachers value assessment practices that differ from the ones to which school funding is tied.

Two years after the Bill and Melinda Gates studies were released, a smaller national survey (n=1500) conducted by the National Education Association (Walker, 2014) showed that testing continued to affect classroom practices in ways that 42% of respondents considered to be

negative. This study shed light on another aspect of teachers' views on accountability; 45% of teachers surveyed indicated that they had considered leaving the profession altogether because of standardized testing. This finding is particularly interesting in light of the fact that standardized testing in the U.S. has existed easily since 1900 and has been standard practice since World War I (Alocer, 2017). This finding begs the question of whether or not the test-related practices and their effects on teachers are the impetus behind this considered exiting of the profession.

In 2016, the Center on Education Policy published the results of another nationwide (n=3,328) survey of teachers (Ferguson, Kober & Rentner, 2016). This study provided the most up-to-date quantitative representation of teachers' perspectives on standardized testing practices. The findings echo those of earlier studies, namely that results from standardized tests affect the way that teachers teach; sixty-eight percent of math teachers and 71 % English language arts teachers changed their instruction based upon 2015 standardized test results. Similarly, many of these teachers expressed a belief that they spend too much time preparing students for standardized tests (62% for state-mandated and 51% for district-mandated tests), and 81% believed that students spend too much time taking mandated tests.

While primarily quantitative, this study included several open response questions; on these questions, the teachers "wrote in almost equal measures about their desire to help and support students and their frustration with an education system that is too focused on testing" (p. 6). This finding indicated that just two years ago, teachers felt that helping students was, at least in some regards, antithetical to helping them pass tests. It would appear that what teachers valued as helpful differed from the help that they were required to offer students in the form of preparing them for tests. According to this study, teachers nationwide still felt pressure to conform to accountability measures. If the teachers generally agreed with the policies enacted,

perhaps this finding would be of no consequence; however, nearly half (46%) of teachers cited state or policies as a major challenge to their ability to teach (and these policies are responses to national legislation) Thus, the disconnect between what teachers believe about how they would like to teach and how they actually do teach still existed in 2016. Even though ESSA lessens some of the most stringent demands of NCLB, accountability and (failure to measure up) is still measured in terms of AYP, and AYP is still primarily measured by end-of-year tests. Thus, the function whose output seems to be testing-related practices disliked by substantial numbers of teachers has, in reality, not changed much.

In contrast to teachers' opinions about their voices affecting state and national policy, many (53%) believed that, at the local and school level, their opinions were taken into account. For this reason, the research I conducted is meant to open conversations at the local level; grassroots discussion is a primary aim of this study.

Qualitative Studies

Introduction. The quantitative studies I reviewed demonstrated that in the sixteen years since NCLB was passed, teachers studied have consistently taken issue with some of the practices associated with mandated testing programs. However, in order to gain a fuller understanding of teacher perspectives, I turn now to the qualitative literature from the same period (2002-2017). Merriam and Tisdell (2016) described the goals of qualitative research as “understanding how people interpret their experiences, how they construct their worlds, and what meaning they attribute to their experiences” (p. 6). If the quantitative studies demonstrated that the NCLB-related phenomena of curricular/pedagogical changes, testing pressures, and teacher discontent exist, these qualitative studies deepened understandings of those phenomena by delving into the detailed meaning making of teachers who experienced them.

Rather than replicate experimental results, these studies reiterated and provided different angles of entry into the problems of practice teachers face under local permutations of NCLB policy. They often did so by providing, in addition to generalized findings, participant quotes. In this way, the qualitative studies dovetail with the quantitative research of the same period to provide vivid recountings of the teachers' experiences. The studies I review contain not only numbers and words but also cognitions and, at times, emotion. In this way, the research promotes multiple truths, which considered in light of larger, previously reviewed studies yields a more complex, multidimensional understanding of post NCLB testing/teaching practices.

Qualitative studies 2003-2007. Hinde (2003), conducted a study of elementary teachers (n=16) in the southwest which she self-described as an interpretivist narrative. With her work, she sought to suss out “why and how teachers believe they have made significant changes in their practice as a result of standards and assessments” (p. 3). Intrinsic to Hinde’s design was the not only the desire to report research *about* teachers but also to include the voices of the teachers themselves in her write up. This she accomplished by threading the words of her participants throughout her findings. Hinde concluded that teachers in the study felt pressured to raise test scores instead of utilize differentiated curriculum and were offended by the practices they were encouraged to pursue; she supported those findings with evocative, image-laden quotes. In describing pressure to help students pass tests, one focus group participant asserted, “I feel like we have a little ax over our heads or whatever. Because there is so much pressure for us to meet our goals so the whole school can get money . . . And you kind of feel like, God, you'd hate to be the one grade level that makes the school not get their money” (p. 11). In answer to Hinde’s primary research aim, she offered the following conclusion: Teachers see the changes to their practice (pressure to raise test scores, competition amongst teachers, curricular pressure) as a

direct result of the legislative mandates of their time. Moreover, they see the emphasis on test results as “sick and wrong” (p. 10). These strong emotional responses to testing-based practices are akin to those I observed in my own teacher participants during the pilot (Richter, 2017a) and in this study.

Shortly after Hinde’s study was published, Craig (2004) conducted a narrative study, which examined “how mandated testing for public reporting purposes influenced educators’ experiences” (p. 1230). In this multi-year study, Craig discovered that accountability measures created a sort of squaring off between what mandated testing required and what the administrator and teachers’ believed was in the best curricular interest of students. She summed up her finding with this comment by the school administrator: “One part of the conflict is that we are going to have to deal with the school’s test scores. The other part of the conflict is we have some reforms we want to do, and they involve application of knowledge rather than accumulation of knowledge . . . We want our curriculum to be full-bodied, deeply textured, rich, relevant” (p. 1244).

This administrator’s comment sheds light on an interesting component of the curricular-narrowing conundrum. Throughout the narrative, it is evident that pressure to improve test scores drove much of the school’s practice. Yet, as a whole, this school community acted collaboratively to meet the demands of the accountability system not by narrowing curricula but by enacting practices that were driven by culturally-relevant pedagogy. The approach worked, though it took several years to bear fruit in the form of increased aggregate and subgroup scores. In the meantime, the school suffered the loss of its principal who succumbed to accountability pressures and resigned. This study is significant in that (1) the educators recognized the pressure

to curtail curricula and chose not to and (2) that the approach worked and (3) the school largely served students (demographically) targeted by NCLB mandates.

In contrast to the high school highlighted by Craig (2004), Agee (2004) focused on the case of a single high school English teacher who found her teaching identity at odds with the required demands of data-driven school. This new teacher, an African American female practicing in New York, desired to implement constructivist, multicultural instructional approaches; yet the environment in which she found herself hindered her progressive teaching inclinations. The teacher believed that the stringent measures enacted by the school stifled her ability to incorporate diverse points of view, to be the multicultural teacher she envisioned herself to be. Craig's study illustrated a tension inherent in test-driven environments. As curricula align to tested material, especially if such curricula are scripted, then implementing multicultural approaches can prove problematic. Multicultural pedagogues attempt to situate curricula in the lives of their students. Using diverse texts and an understanding of students' lives, multicultural educators create specific, unique, class-specific curricula. Certainly, the teachers in my pilot study (Richter, 2017a) indicated that they wished for the ability to create such curricula and, specifically, to let their students (the majority of whom were African American) see themselves represented in texts, but the mandated curriculum and brisk pacing prevented them from inhabiting the multicultural practices they ideologically embraced. An African American, eighth grade English teacher clarified the matter thus: This teacher's scripted curricula utilized primary sources (a letter written by Frederick Douglass, an image of a slave). And while one of her students specifically asked to know more about the author or about the conditions that led to the pictured slave, my study participant believed she was not allowed to engage students beyond ways her curricula dictated. Because this teacher perceived pressure to

shut down inquiry and to stick to her mandated script, she answered her student, “No. I can’t teach that” (Richter, 2017b, p. 143).

While Agee’s study demonstrated that in some contexts, incorporating multicultural texts and viewpoints while preparing for standardized tests can be a daunting task, Bolgatz (2006), offered a different perspective on the matter. In another New York-based study, Bolgatz used primary source documents in order to help students prepare for end-of-year standardized tests in social studies. Bolgatz utilized the texts to engage students in test prep by aligning tested skills with in-class tasks around these texts, which centered on issues of race. But Bolgatz didn’t stop with test prep; using the texts, this teacher-researcher started conversations about issues of racial disparity. Bolgatz’s case study, while promising in itself, highlighted the complex, context-specific nature of the teaching/testing conundrum. In Bolgatz’ study, the social studies test was not attached to high stakes for either student or teachers. I suspect this is one reason that Bolgatz had the autonomy to engage students in both testing-related and discussion-based, multicultural practices.

In an essay addressing testing and teaching in high-stakes testing environments, Dufy (2007) offered further insight into the circumstances that foster this double-duty type teaching: School leaders’ views of literacy. In schools where leaders recognize literacy as more than test scores and value teacher autonomy, school leaders understand that, “The problem is that rules, regulations, and pressures designed to ensure high test scores replace ‘best practices’ with more drill and practice, and replace professional teaching with technical compliance” (p. 9). Perhaps this view of literacy could embolden school leaders to give teachers like Agee (2004) and the teachers I interviewed the license to both teach to the test *and* teach in culturally-relevant ways.

In one of the largest, urban qualitative studies to date (n=200 interviews) Croco and Costigan (2007) highlighted the discrepancy between new teachers' reasons for entering the teaching profession and the constricted teaching spaces they felt forced to inhabit because of testing-related practices. Over a five-year period straddling NCLB's passage, the authors documented curricular narrowing, teaching to the test, and test-related pressures in their New York City based schools. Of consequence to my study is one of Croco and Costigan's main findings. The authors stated of their teacher-participants, "They believe that scripted lessons and mandated curriculum not only de-professionalized their work but also depersonalized the human connections nurtured by more student-centered curriculum and pedagogy" (p. 521). In one of the largest, longest, and most recent qualitative studies to date, teachers from two separate cohorts (English and social studies) reacted strongly against testing-related practices, which they labelled as de-professionalizing and depersonalizing. Such disparity between beliefs and actions is precisely the type of disconnect my study addresses. Given the meticulousness of the authors' methods, the length and timeframe of the study, and the number of interviews and participants, saturation of analysis seems likely. Thus, such strong reactions to curricular-practices are worthy of further consideration.

Before moving on from this study, I want to reiterate the way in which Croco and Costigan supported their findings with particularly poignant participant quotes. These researchers' stated goal was to include teacher voices in order to "offer another perspective on the statistical data provided by economists, sociologists, and policy makers concerning the effects of NCLB" (p. 3). Considering the finding that nationwide, many teachers believe their opinions factor in little in the decisions that dictate their practice (Ferguson, Kober, & Rentner,

2016), research like Croco and Costigan's study (and mine) that serves to center teacher voices is sorely needed.

The extent to which participant voices are included in data write-up varies greatly amongst the research I have reviewed. While thick descriptions are considered standard in qualitative texts, the extent to which Croco and Costigan incorporated them added another layer of complexity to their findings. Participant quotes like this one, "I have to cut out certain cooperative activities because they're time consuming. It definitely affects my teaching. It's always in the back of your mind . . . certain topics I would expand on, especially if they were relevant to the kids or of particular interest to me, but I don't because they're not tested on it" (p. 521), vividly illuminated the lived experiences, the exact ways, in which testing can drive teachers to narrow curricula. In this case, the teacher expressed a tradeoff between student interest and topics actually taught. This nuanced, up-close view of curricular narrowing and teacher responses foreshadows the in-depth narrative research I conducted in New Orleans.

Croco and Costigan (2007) focused on the perspectives of teachers themselves in English and social studies in an urban environment. In contrast, Willis (2007) considered how non-tested curricula are affected when time is taken from them to teach tested subjects. Willis explored this issue in a low-performing, rural, California elementary school, which served a mix of Latinx, African American, and White students. In regards to social studies curriculum at the school he studied, Willis concluded that "the scope and sequence of the social studies curriculum and opportunities that would deepen students' understanding of history are being squeezed from the enacted curriculum" (p. 1980). Willis's study is of particular interest because it detailed concrete effects of curricular narrowing in non-tested subjects. Trickle-across-the curriculum effects are logical extensions of curricular narrowing, but Willis addressed them in specific, concrete terms.

Perhaps the most extensive study of curriculum-specific effects of the accountability movement was published five years after NCLB passed. Au (2007), synthesized the results of 49 qualitative studies; he utilized a particular form of metasynthesis, template analysis, in which he constructed codes to analyze the studies and present the results as a whole. Au employed metasynthesis in order to look for connections between high-stakes testing and curricular changes. The 49 studies Au analyzed spanned 19 states from all regions of the United States. (No Louisiana studies were included in the analysis, presumably because they didn't exist).

Au's study is particularly useful for categorizing the different ways in which curricula have morphed in order to align to high-stakes tests. Those changes have been documented across three main categories: Content, knowledge form, and pedagogy. In 69.4% of the studies analyzed, Au found that curricular content is contracted (narrowed) so that non-tested subjects are given less instructional time. Concurrent with this content contraction, Au documented a phenomena he called curricular expansion. In 28.6% of studies analyzed, tested content bled into non-tested subjects (e.g. teaching tested writing skills in social studies). In regards to knowledge form, Au analyzed whether or not teachers shifted the form tested knowledge took. In 49% of the studies, Au coded examples of knowledge fracturing, that is "knowledge forms being fragmented and isolated into discrete test-driven bits" (p. 260). In the final category coded, pedagogy, Au documented a test-related shift, citing that in 65.3% of studies reviewed, in response to high-stakes testing, educators began to implement more teacher-centric approaches such as lecture and direct transmission of test-related facts. Au asserted that, considered in conjunction with each other, the substantial changes in content, knowledge form, and pedagogy are evidence that high-stakes testing is form of testing-incited curricular control.

Au's metasynthesis cemented the notion that high-stakes standardized tests significantly affected curriculum. Valli and Buese's (2007) five-year, longitudinal, elementary-level study honed in on the effects that these changes had on day-to-day the tasks that teachers performed. Conducting 16 focus groups with 84 teachers, Valli and Buese ascertained that policy changes during the early 2000s inspired quick, breadth-over-depth curricular pacing and content/test alignment. In low-performing, low-income schools, these effects were exaggerated and much of the inspiration for curricula came from end-of-year tests. In addition to these findings, Valli and Buese noted that in the immediate years post-NCLB, teachers took on additional tasks of data managing and, as AYP pressures mounted, began tutoring targeted student groups both prior to and after school. This study did more than just document the effects of increased tasks; the authors specifically linked the increases in teacher tasks to state, district and national policy changes, and they noted the year that many of those tasks came into being . . . 2003. Thus, this study clearly nested increased teacher workloads within an era of accountability policy implementation.

The authors concluded their analysis with an assertion that, if true, merits a substantial discussion in its own right. Valli and Buese concluded that the extent to which the tasks teachers undertook changed so much after NCLB that these changes represented a difference not only in degree but also in kind. According to these authors, accountability-driven mandates changed the very nature of these teachers' jobs; as a result of national policy, their role in the local system had fundamentally shifted. Specifically, teaching morphed from being primarily a self-regulated instructional job to institutional, top-down, externally mediated one.

Valli and Buese's hypothesis is relevant to this study because if policy changes are driving the way teachers in New Orleans conduct instruction, and if teachers are dissatisfied with

those changes, the question of how teachers conceptualize their roles versus the current ways those roles are embodied represents another strand of a cognitive disconnect that is worthy of exploration. Data from the pilot (Richter, 2017a) indicated that this disconnect may be locally present. Concerns from my participants that fast-paced curriculum, scripted curriculum, and departmental meetings centering around data mirror this study's results. And, like the teachers studied by Valli and Buese, each study participant indicated a concern that the relational aspects of their jobs (which they found highly rewarding) were impinged upon by the instructional and institutional tasks they were required to perform (Richter, 2017a).

Qualitative studies 2008-present. Arguably, the studies reviewed thus far represent expected changes in the first few years of the NCLB era. In response to the new legislations, the trajectory of the ship shifted. Whether or not these changes represent a righting of the ship that resulted in measurable gains for targeted students is unclear. What is evident is that within five years of its passage, NCLB invoked and exacerbated test-related practices. While some teachers appreciated the structure and accountability that the policy changes provided, others found that testing practices did indeed undermine their notions of what it meant to teach.

If the literature indicated that these first few years were just an adjustment period and that teachers had, more or less, reconciled themselves to the new normal, perhaps this study would be unwarranted. However, researchers in this second wave continued to document the phenomena of test-related curricular changes and teacher discontent to those changes. I have located fewer studies conducted during the years 2008-2017. It is my belief that the slowdown in documentation is not due to change in the opposite direction (widening curricula, teacher satisfaction with testing practices). If this were the case, I would have likely unearthed studies

documenting these changes. More likely, the earlier research had done its job well. Fewer researchers seemed to be addressing how curriculum and pedagogy changed because of NCLB. A common theme of this second wave is a collective concern over how teachers were handling these changes. They do document the changes, but *that* the changes occurred is, at times, peripheral to the discussion. The researchers in this wave were attempting to bring the teachers into the center of the stage.

In addition to documenting frustration and stress related to curricular changes and testing pressures, Costigan (2008) addressed the questions of how teachers working under stringent accountability regimes might successfully navigate testing mandates and still implement practices in which they believed. Craig's (2004) teachers wanted to implement full-bodied curriculum but came under fire when the first year they did so, test scores dropped. Agee (2004) intimated that she found it near impossible to teach in culturally-relevant ways. Bolgatz (2006) infused test-prep with primary source texts and engaged in discussion-based teaching. Focusing on the experiences of 12 urban English teachers in New York, Costigan unearthed a single success story of a ninth grade teacher, Michael, who untethered himself from the data-driven demands in a low-income, urban school.

After teaching the mandated curriculum for two years, Michael and a cohort of other experienced teachers informed their administration that "there is no way in hell we're doing the scripted curriculum!" (p. 91). This pronouncement procured Michael a certain amount autonomy over his instructional choices; he was able to assert himself thus after gaining the respect of his administration and teaching the scripted curriculum for two years. Even though Michael reported satisfaction in being able to plan the lessons he deemed appropriate, he still regarded the constant supervisions, bulletin board mandates, and room arrangement requirements with disdain. Perhaps

the lesson from this success story might be “Do your mandated curriculum time, then take a stand, or “Choose your battles.” The authors merely report. They do not speculate that this success could be easily replicated.

Michael’s approach was simple, direct, and perhaps a bit rebellious. The three local teachers I interviewed (Richter, 2017a) certainly indicated the ways in which they would navigate around testing-practices would be less frontal. One teacher described looking for loopholes in the curriculum that would allow her to follow inquiry-driven lines of instruction. This same teacher indicated that change might occur as she learned to trust her teaching sensibilities and find ways to lessen her fear of the test (Richter, 2017a; Richter, 2017b). Similarly, the high school English teacher I interviewed located particular places in her curriculum where she could implement pedagogies in which she believed. This teacher, arguably under less pressure than her elementary and middle school counterparts, found ways to help her students understand some of the whys behind her curricular choices even while formatting instructional materials to match the test.

Instead of focusing on the actions of teachers who took testing-related practices to task, Newberg-Long (2010) contributed to this second wave of post-NCLB qualitative research by conducting an in-depth, phenomenological exploration of the teachers’ perspectives on test-mediated curricular changes. As the only phenomenological study (n=3) I’ve uncovered to date, Newberg-Long’s study sought to extensively document the beliefs and perspectives of urban, elementary school teachers in regards to curricular narrowing, roles and responsibilities, and accountability expectations.

Newberg-Long reaffirmed the findings of first-wave research: Her participants described feeling pressure to teach to the test (greater than 10 years prior to the study) and deliver narrow,

fast-paced curricula. They also reported a conflict between their beliefs about pedagogy and district mandates. If the study had stopped with simply reporting these findings, it would have been significant as identifying another pocket of teachers who describe a distasteful experience with standardized testing in the age of accountability. She did not. Similar to Croco and Costigan (2007), Newberg-Long believed that including teachers' voices in her write-up was central to her research. Newberg-Long's analytical task was to deduce and present the essence of the phenomenon under study. What makes this study significant and precedent-setting in this body of literature is not so much the essence of these educators' experience, which she does. It is the *way* in which Newberg-Long chose to represent that essence — through a series of poetic narratives — that sets this study apart from its predecessors.

In describing her rationale for this methodological choice, Newberg-Long argued that writing poetic narrative helped her distill the phenomenological essence by “allowing for selection of words and phrases most meaningful to the interpretation” (p. 78). In other words, the art form aided her analysis. Furthermore, she reasoned, “poetry is welcome in diverse settings such as poetry bars, policy making settings, and the mass media” (p. 79). Newberg-Long believed that writing her findings as poetry could further the impact of her findings by allowing them to be publicly published (i.e. performed).

Though Newberg-Long did not frame her research as such, she used an arts-informed approach to render her participants' experience in evocative ways (Merriam & Tisdell, 2016). Each finding was tied to a narrative piece comprised of participants' voices. The choice to represent data thus was intentional on the part of the researcher as she chose to “represent emotions and feelings experienced by teachers, as well as facts” (p. 78). I call attention to the methodology of this study because, to my knowledge, it is the first of its kind in the

accountability literature. Because this type of data representation may be unfamiliar to readers, I provide an example of a poetic narrative from a study participant, Liz. The poem occurs under the heading “Focus on Tested Subjects: Reading, Writing, Math” (p. 92).

Writing is not used for AYP
So reading and math are our focus
We are looking at a 3 hour literacy block for next year
Then math for 90 minutes

This much time for reading
This much time for math
This much time for writing
Because it’s all assessed
(We teach) Test Taking as a genre
You’re going to think I teach to the test
But I really don’t
I need to do some things so my kids have some experience with this
CSAP (asks for) summaries
They didn’t know the word “summarize”
We realized we needed to give them this vocabulary

Similar to Newberg-Long, I chose to incorporate artistic process in my pilot study “Now That Testing is Over” (Richter, 2017a). I made this decision for reasons similar to Newberg-Long; namely, it was my belief that the data came imbued with emotion as well as fact. With the

musical narratives I wrote, I restoried these multiple dimensions of the data. That is, I tried to musically and lyrically portray both the facts and emotions of my study participants' stories. I replicated that approach in this study.

This ability of art as research to render and evoke emotion is one that arts-based researchers accept and exploit for its research potential (Leavy, 2013; Leavy, 2015). Like Newberg-Long, I also rendered my findings in the first-person voice of my study participants. Harkening back to Rapp's (2002) call, I desired to write research that was not only *about* teachers but that centered their words in the process. Finally, like Newberg-Long, I saw the performative potential of musical narratives as a significant way to potentially extend the reach of the study. Performances of the narratives in small settings (an educational conference, a doctoral level research class, for the study participants themselves) confirmed that this type of data rendering evoked both intellectual and emotional responses leading to discussions of the research issue. This study furthered this performative approach by singing the stories of the study participants in the community at large.

Moving in this second wave of post-NCLB research Santoro (2011) mostly ceased addressing the question of whether or not teaching in an accountability-driven system changed teacher practice. In this second wave of qualitative exploration, researchers deepened understanding of test-related phenomena, posited rationales, and in some cases offered a heuristic through which to view the phenomena. It is this task of interpretation, an embryonic stage of theory building, to which Santoro attended via a Virginia-based case study and philosophical inquiry.

Santoro's particular contribution to the accountability conversation was not so much the idea that teachers are simply dissatisfied, stressed out, frustrated, or fearful of falling short on

accountability measures. Rather, when teachers such as Santoro's study participant could no longer access the moral rewards of their work, Santoro posited that they experience a type of moral depression. Santoro's participant, Stephanie, fought and fell prey to this moral malaise when she attempted to operate in an increasingly test-driven climate. According to Santoro, "What had been hallmarks of good teaching for Stephanie — connecting student learning with their experiences, helping them learn to think in ways that will transfer to success in higher-order analysis and their everyday needs, and maintaining creativity in her work and her students' problem-solving — was being jettisoned by the exigencies of passing the test and making AYP" (p. 16). Eventually, after seeing the test-preparation materials she had been asked to create for review *become* the curriculum, Stephanie left her school. She did so on moral grounds.

Santoro was careful to avoid framing this this type of leaving, which she typified as conscientious objection, as burnout. Stephanie indicated that she was not overwhelmed by her job's tasks or the needs her low-income students. Instead, Stephanie's choice to leave her high-poverty school, like 12 other teachers Santoro studied (Santoro, 2011a), stemmed from being unable to square her moral beliefs about teaching with the actions imposed on her by her particular school system. In an effort to create a common language around the types of events described in this case study, Santoro gave this set of behaviors a name: Demoralization in the pursuit of good work.

While I certainly did not plumb the depths of the local teachers' perspectives on their beliefs about the morality of teaching, I do believe the seeds for this type of further exploration were present in the pilot data. In particular, I am reminded of a focus group participant's comment: "I was born to do this (teach), so I'm gonna do right by my kids" (S. Smith, personal communication, March, 14, 2017). This study participant described her teaching in terms of

making ethical judgement calls. She thought of her practice as having right and wrong expressions, thus implying that her work had moral rewards, which could be both accessed or denied.

In 2016, Augustine (2016) published narrative research recounting the perspectives of 12 Indiana elementary, middle, and high school teachers regarding the introduction of high-stakes testing into their classrooms. This study uniquely explored the changes in classroom dynamics that occurred as high-stakes testing practices took root. Augustine's inclusion criteria assured that she would be gathering data from only veteran teachers who had each been practicing nearly continuously for twenty years. Her participants had all been teaching several years prior to the passage of NCLB and continued to practice during the decade after it was passed. Given that this study is a narrative, the extended tenure of each of the teachers situated the researcher well to narrate plot of changing classroom dynamics in response to the inciting incident of accountability measures.

In the wake of that inciting incident, her teachers shared a common narrative thread: Their practices had change drastically in the years following NCLB, and those changes persisted into the present (time of data collection). Their stories echoed the experiences of other teachers researched post-NCLB implementation. Those experiences included frustration with increased time spent preparing students for standardized testing. These participants noted that the time they spent has "continued to increase as more tests were added to the battery" (p. 114). In addition to frustration, these teachers keyed in on stress that they felt to prepare students for standardized tests *in* standardized ways. The author named this the "lockstep" lesson and timetable approach to curriculum planning (p. 115), and noted it as major pedagogical change. In response to this

change, teachers indicated that lockstep approach hampered their creativity and reduced their joy.

Like Newberg-Long (2010), Augustine paid attention to more than just the facts of the data she gathered. When Augustine wrote the following, “It became obvious from these teachers’ remarks and their voice inflection as they answered my questions that their frustration level was elevated and they felt mentally drained from administering the exams” (p. 111), she was calling upon and interpreting non-verbal cues. Knowing that Augustine member-checked her research, I assume her participants validated her interpretation.

My recognition that qualitative interview data can be “imbued with emotion” (Richter, 2017a) mirrors Augustine’s. Whereas Augustine referenced teacher nonverbals in passing, I attempted to interpret those cues and restory them as part of the final product. I did so in hopes that listeners could possibly connect with participants’ emotions on a visceral level. I continued to that approach with this study; that is, I attuned myself to the emotional, nonverbal cues embedded in interview data and rendered those aspects of the data as accurately as I could in the songs I wrote.

Arroyo-Romano (2016) also tuned into the emotions of participants while recounting the experiences of two high school teachers practicing in a rural, southwestern state. Describing the effects of testing, one teacher stated, “You can feel yourself feeling the pressure and clenching your jaw and snapping at people maybe a little more” (p. 64). The other participant recounted, “. . . obviously stress is on myself, stress is on my students . . . As far as the school goes, it is very stressful, and especially this time of the year because the test is next week” (p. 64). Arroyo-Romano examined this stress and the pedagogical choices that teachers made through an ethical lens. Drawing upon virtue ethics (Campbell, 2003), the ethic of care (Noddings, 1984), and

Paulo Freire's ethic of justice (Freire, 2004), Arroyo-Romano deduced that teachers in this test-driven context were navigating pedagogical ethical dilemmas. According to Arroyo-Romano's conceptual framework, the ethical teacher considers the needs of students and attempts to meet them. One way ethical teacher meets students' needs is by planning and implementing quality instruction. In this ethical framework, these tasks are elevated to the level of duty.

The teachers in Arroyo-Romano's study also viewed their duties as "helping the students graduate and facilitating their vision for a future" (p. 69). The ethical dilemma arose when the teachers tried to balance those duties with the policy-imposed tasks of preparing students for high-stakes exams. Doing right by their students meant walking a tightrope between what these teachers considered important and what policy demanded. After all, part of doing right by students needed to pass standardized test in order to be promoted to the next grade level, graduate, gain college admission. Therefore, unlike Santoro's (2011) conscientious objectors who existed the profession, these two teachers "compl[ied] with the norms and mandates dictated by the state and the district and [made] educated ethical decisions based on their perceived responsibilities and knowledge of their ethical values" (p. 69). In other words, they made pedagogical choices that they hoped would meet both the demands of their personal ethics systems and the dictates of the accountability system.

Santoro (2011) and Augustine (2016) focused their research on the ethical and moral aspects of the teaching/testing conundrum. Their research sheds light on how dissonance may arise (as the result of moral beliefs clashing with required duties). I have seen this connection between morality and cognitive dissonance fleshed out in the psychological literature as well (Graham, 2007). As a result of reading Santoro's and Augustine's research, I incorporated questions into my interview protocols that pertain to these dimensions of the problem of practice.

No matter which angle I explore the problem from, I return to questions of values (what teachers value, how those values would ideally be expressed in classrooms). These studies provide a language with which to address those issues in my own research.

New Orleans Context

Thus far in this literature review, I have reported on and analyzed studies regarding the pedagogical effects of nationally-mandated testing during the years 2002-2017. From the studies published during this period, I have also garnered information about how various groups of teachers in differing contexts have responded to those effects. When appropriate, I have demonstrated how different studies either build upon or differentiate findings from their predecessors. And finally, I have drawn connections between the literature at large and my own research. At times, I have highlighted similar findings; on other occasions, I have noted methodological similarities, some of which inform how I conceptualized and conducted this arts-based research study.

What I have intentionally neglected to review until now are studies that locally address issues pertinent to my study. However, I now turn my attention to the context in which this study takes place: New Orleans. In the wake of Hurricane Katrina's citywide destruction, the school system underwent a massive restructuring (Garda, 2010). Seeing the physical destruction of the city as the ash from which a new, better system could be reborn, the state fired all of the public school teachers (Bell-Weixler, Barrett & Harris, 2017). This razing (and subsequent rebuilding) of the New Orleans school system happened three years after NCLB was signed into law. Thus post-Katrina schools were concurrently shaped by both physical and legislative forces active at the time. Ever since Katrina, the city has been transitioning from a public to a primarily charter run system (Bell-Weixler, Barrett & Harris, 2017)

This move toward charters was and is driven by the accountability movement. According to Holley-Walker (2007), “The New Orleans public schools after Hurricane Katrina are a case study that demonstrates that accountability measures, like NCLB, may lead to the decline of the public school system and replace that system with privately run, autonomous charter schools” (p. 125). Holley-Walker’s prediction has borne out in New Orleans, where it appears highly likely that at some point in 2018 the city will be the first all-charter system in the country (Vanicore, 2016).

In local discussions of school reform, accountability is a recurring motif. In 2017, The Louisiana Association of Public Charter Schools (LAPCS) stated that, “Charter schools are held accountable for improved student achievement and closed if performance is consistently low” (“About Charter Schools,” 2017). The performance to which LAPCS refers is a school’s performance score (SPS), which is primarily comprised of student performance on state-mandated, standardized assessments (“Accountability, School Performance Scores,” 2015). As a whole in 2017, Louisiana school performance scores rose (Nobles III, 2017); New Orleans school performance scores did not follow suit (Jewson, 2017). For three local schools, the stakes attached to standardized testing were very recently actualized as their charters were not renewed (Nobles III, 2017). New Orleans has been, and remains a city where tests are decidedly high-stakes.

That the current iteration of the New Orleans school system has been shaped by the accountability movement is clear, but what effects have this attention to accountability wrought upon the ways in which curricula are conceived and delivered in New Orleans? With so much attention given to the ends, which studies focus on the means, and the effects those means have had on local educators?

Dixson, Buras and Jeffers (2015) claimed that post-Katrina reform efforts cut short plans to involve the local community in curricular decisions making. In support of this assertion, they referenced the disenfranchisement of the Frederick Douglas Community Coalition located in the upper ninth ward. As part of the local Pre-Katrina effort to provide equitable education to all students, the 2005 coalition planned to “more centrally involve the community in decision making over curriculum and educational policy” (p. 291). The authors did not state what specific curricula would have emerged from this effort; they did, however, argue that an influx of non-local, charter administrators and educators stymied the efforts.

Bell-Weixler, Harris and Barrett (2017) studied the effects that New Orleans’ massive reform had upon teachers (n=323) who worked in the city both prior to and after Hurricane Katrina. The majority of teachers surveyed (60%) indicated “their schools and their lives as teachers are worse under the reforms” (p. 20). The authors posited that in conjunction with longer hours and decreased job security, higher-stakes teacher evaluations contributed to this perception. This study is useful because it situated teacher dissatisfaction as the confluence of multiple variables. What it did not do is weight these variables; nor did it attempt to suss out how they might interact with each other. However, longer hours, decreased job security, and higher-stakes evaluations are all related to accountability. The movement continues to affect school culture in ways that local teachers find dissatisfactory. Bell-Weixler, Harris and Bennett asked the obvious question of whether or not this dissatisfaction contributes to teacher turnover. Certainly, the teachers I interviewed were dissatisfied with how accountability affected their day-to-day practices and intimated that had they the power to do so, they would not let testing dictate their practice to the extents that it did (Richter, 2017a).

In an attempt to provide a “direct eyewitness account of the successes and failures of the city’s new direction in public education,” (p. 463) Ciolino, Kirylo, Mirón, and Frazier (2014) interviewed 20 elementary, middle, and secondary teachers. The teachers studied named several issues that they encountered in the post-Katrina, charter-reform era. Some of these issues are peripheral to this study — lack of quality school access for certain populations, the negative effects of “school shopping” on school’s student populations — but one is directly related. Teachers interviewed expressed concern over their lack of job security, which they perceived as directly related to how well their students performed on standardized tests. One teacher stated, “. . . you don’t have a contract [and] they could legally come in and say, ‘We don’t need you anymore . . . If you’re not bringing the test scores up, you could be let go’ ” (p. 465). If bringing up test scores is tantamount to job security, it is reasonable to assume that teachers would modify curriculum and methods in order to meet this goal. As the review of national literature demonstrated, such test-driven, curricular changes have been documented throughout the country. What remains to be seen is how curricula and instruction may have morphed in New Orleans, where to date, post-Katrina, over 25 schools have been closed for failing to make the grade (Bell-Weixler, Harris, & Barrett, 2017).

The only New-Orleans study examining curriculum and pedagogy is a recently published ethnography (Sondel, 2016). Sondel identified the same local literature gap that this study helps address. In her study rationale, she reasoned

Significant theoretical and empirical work has analyzed how post-Katrina reforms diminish community participation in policy decision-making processes, incentivize push-out, and otherwise fail to meet the needs of low-performing students (Bordelon, 2010; Buras, 2014; Lipman, 2011; Wolf, 2011). Yet underexplored in the literature is the

impact of these policies on curriculum and instruction. This article aims to address that gap. (p. 172)

Sondel's ethnography of two New Orleans Charters schools complements Dixon, Buras and Jeffers' (2015) work by concretely exploring how reform played out curricularly in two local charters. I closely examine this study, in particular, because it is recent, it is local, and it addresses similar themes to both my pilot and dissertation research.

Sondel spent six months in two schools observing and interviewing study participants. Unlike the majority of studies reviewed, where interview seemed to serve as the main data-gathering tool, Sondel elicited much of her data from observing teachers instruct in real time. Those observations, she analyzed through the lens of culturally-relevant pedagogy (Ladson-Billings, 1995; Ladson-Billings, 2009). The framework encompasses six aspects of pedagogy, and Sondel explained her findings in light of each.

In regards to the first aspect, academic rigor, Sondel described the starting point for curricular planning as a type of backwards mapping in which school staff poured over the prior years' test results in order to determine which standards should take top priority in the upcoming year's scope and sequence. In other words, the curriculum was mapped to match the tested standards. Skills and content deemed peripheral to the test were relegated to be taught after the test or not at all. What Sondel described is a type of school-wide curricular narrowing taught by a sort of lockstep lesson plan to which Augustine (2016) alluded. That lesson plan followed the following identical five step format: "1. The teacher introduced the lesson objective. 2. The teacher modeled the application of the objective. 3. The teacher practiced the objective with the students. 4. The teacher had students practice independently. 5. The teacher assessed 'student learning' for 'mastery' " (p. 177). Throughout this process, Sondel noted few references to the

intrinsic value of learning. (Such practices would be culturally relevant.) Instead, material was introduced through statements referencing its importance to passing standardized tests and was taught in decontextualized ways (e.g. lifting lines from poems to teach about metaphors but not reading the entire poem).

In regards to Ladson-Billings' second pillar of culturally-relevant pedagogy, cultural competence, Sondel documented little to no situating of students' cultures into the classroom (note: 93% and 98% of students in these two schools were African American). Instead of situating pedagogy within students' cultures, curriculum in these two schools relegated student cultures to peripheral or extracurricular activities. While students might attend an afternoon workshop hosted by Mardi Gras Indians or celebrate Black History Month, African American texts were not incorporated into the curriculum. While Sondel noted a few instances of teachers incorporating local culture into their actual lessons, she primarily reported upon utilizing culture as a form of control. In one instance, students were able to attend cultural events as reward for compliant behavior. In another, students who misbehaved were required to copy lines of Martin Luther King's "I Have a Dream" speech. In the first example, the hidden curriculum dictated that cultural events are unimportant and simply add-ons to *real* education. The second example situated cultural texts as a vehicle for punishment.

Pedagogy that embodies the third component of Ladson-Billings' framework, critical consciousness, encourages students to critically reflect upon the values and norms of the social institutions with which they interact. Sondel argued that with little to no connections to students' lives or cultures, the curriculum of both schools lacked this aspect of culturally-relevant pedagogy. Further, the strict behavior policies she observed, walking in silence in straight lines, learning in silence, students keeping their bodies in specific postures, all represented a hidden

curriculum of compliance both in the classroom and out. She concluded that the stringent behavior monitoring via clipboards and apps combined created a culture that failed to foster critical consciousness.

Sondel also questioned the premise upon which the practices she observed was built, namely that higher test scores would lead to college success and a road out of poverty. Such a viewpoint, she argued, was one-dimensional and ignored other components of college readiness such as “a strong racial identity, critical analysis, skills of collaboration, or opportunities to follow one’s own lines of inquiry” (p. 184). Her final assessment of the curricular norms of two New Orleans charter schools suggests that, in cases where standardized testing drives instruction, not only are curricula narrowed, testing-based pedagogy “shifts the purpose of schooling toward the production of assessment data” (p. 185).

The enormity of Sondel’s assertion cannot be overstated. Of course, her statement has not been summarily proven, but it did inform the way I approached teachers whose perspectives I studied in this dissertation research. While I did not specifically ask teachers what they thought the purpose of education is, I did ask them how they defined success for their students and how they thought their institutions defined success. For all teachers, the two answers radically differed.

Regardless of whether or not standardized tests are accurate measures of student achievement, regardless of whether or not aligning curriculum to tests is an appropriate practice or not, if teachers believe that what they are being asked to do is inappropriate they may, like Santoro’s (2011) teachers, consciously object to the methods and materials they must produce. In studying — specifically, locally, deeply — the perspectives of teachers who practice in New Orleans, I am picking up where Sondel left off. Sondel said that, “a full discussion of the lack of

teacher agency falls outside the scope of this paper” (p. 184). What was peripheral but noted in Sondel’s study (teacher concerns over losing their jobs, fear of inadequately preparing students, practicing in ways that they wouldn’t normally choose) is central to mine. In other words, Sondel locally documented the research problem I have explored. My own pilot began deepening understanding of that problem. My dissertation continues to address that problem and contributes to both the local and national literature as another wave of educational reform (ESSA) crests over the Crescent City.

Philosophical Statement

“We see things not as they are but as we are.”

— Anonymous

Introduction

The above quote, attributed to such diverse sources as the Talmud, Anaïs Nin, and even self-help author Steven Covey, encapsulates what I’ve come to believe about the nature of truth — what it is, how we discern and disseminate it to others. For all researchers, the ways in which they view the world inform the research process. Even the topics that researchers deem relevant are mediated by values, interpretations, and personal histories of the researchers involved. While researchers may seek objectivity, it is my belief that worldviews are so entrenched that mitigating them is the best researchers can hope to do.

Human beings, researchers included, are meaning makers. As humans experience phenomena, we do so not as a “buzzing confusion of indistinct and unstructured perceptual elements, but [as] a world that appears as meaningful” (Polkinghorne, 1991, p. 135). The meanings that I co-created with my participants rest upon philosophical presuppositions (mine, theirs). They exist in the dialogues we had through the course of this study, which is built upon a

particular philosophical substructure. Communicating that foundation is part of the due diligence of any researcher, and it is to that task that I now turn my attention.

This study, an arts-based, musical narrative, is undergirded by several philosophical assumptions. As research methodologies go, arts-based research is relatively new. Arts-based researchers incorporate the creative arts into the research process. By melding art and research, arts-based researchers create emotionally-evocative, empathy-promoting works which are imbued with the thumbprint of the researcher (Barone & Eisner, 1997; Leavy, 2015). In addition, arts-based researchers embrace intuitive, non-verbal, non-cognitive modes of knowing (Barone & Eisner, 2012; Leavy, 2015; Rolling, 2013). Several authors (Chilton, Gerber, & Scotti, 2015; Leavy, 2013; Rolling, 2013) have described a need for paradigms that adequately explain inclusion of these ways of knowing in research. Leavy (2013) likened the current state of arts-based research paradigm building to the work that qualitative researchers and theorists undertook when that methodology was emerging. Leavy posited, “If qualitative research has systematically challenged the foundations upon which positivist, quantitative research is built (the subject/object and rational/emotional dichotomies), the explosion of arts-based research over the past two decades has pushed this negotiation even further” (Leavy, 2013, Chapter 3, Arts-based research section, para. 1).

Leavy’s framing of arts-based-research paradigms as extending rather than opposing qualitative paradigms is key. It is not that arts-based researchers reject paradigms upon which many qualitative researchers build their works. In fact, arts-based researchers may engage with research agendas through paradigms commonly associated with qualitative methodologies. Constructivist, interpretivist, and/or transformative paradigms can inform their work. Arts-based paradigm exists in conjunction with, not in opposition to these other worldviews. Such an

addition can necessarily complicate the paradigm conversation or yield a study informed by multiple paradigms. In the case of the study under discussion, three paradigms — constructivist, postmodern, and aesthetic intersubjective — undergird the research.

Constructivism

This study is primarily concerned with a certain type of reality: The pedagogical narratives that teachers construct through their social interactions in their particular teaching contexts. It is the teachers' socially and contextually-derived stories and beliefs that matter, the mapping of their experiences with and beliefs about test-related phenomena that occur in their classrooms and schools. My acceptance of the teachers' reality as truth places this study squarely in a constructivist camp. Constructivist notions of truth — as multiple, context-bound realities (Merriam & Tisdell, 2016) — informed how I elicited, interpreted, and rendered participant data into a narrative whole. In this study, participant realities *are* reality. I trust that I spoke with reliable narrators, that teachers indeed told me their truths. I am not arguing that these truths are universal to all teachers; teachers' truths are neither objective nor generalizable. They are, nevertheless, the truths that this study explores.

As I attempted to restructure these participants' truths via a narrative song cycle, I recognize that I reconstructed a reality out of shared researcher-participant intersubjectivity. Like any narrative inquirer, I am narrating these participants' truths. The truths that I am co-creating as a human research instrument (replete with my own socially-constructed reality) necessitate a set of practices that attempted to ensure, as much as possible, that I made meaning with data in ways teachers deem trustworthy.

Fundamental to accessing participants' truth is an understanding that, as Clandinin and Connelly (2000) stated, "Narrative research is an experience about experience. It is people in

relation studying people in relation” (p. 189). In this research, where participants are few and the subject is sensitive, participant-researcher trust was absolutely vital in order to assume data fidelity. I attempted to garner this trust by engaging in practices that protect my participants’ identities. I also attempted to foster a researcher-participant relationship that was dialogic, close, and as egalitarian as possible. As equal, joint creators of the narratives I performed, my participants had final say over the products I created.

Postmodernism

This study, while guided by a social constructivist perspective, also rests upon the postmodern premises. The constructivist paradigm mediates many of the ontological assumptions that undergird this study. However, epistemological considerations are shored up by postmodernism’s embrace of non-rationality (Merriam & Tisdell, 2016). Embracing non-rationality does not mean that I have abandoned rational approaches to research. It means that rational and intuitive approaches to meaning making are present in the methodology, particularly when I analyzed data and wrote the songs. It is not that the choices I made when setting data to song cannot be rationally justified. They can. However, when I wrote the music, I was engaged in a state of flow (Csikszentmihalyi, 1990). I was completely focused on the tasks at hand, at times making musical choices in the moment based on whether or not the decision felt correct. While I employed a systematic set of practices, allowing for and even exploiting this artistic part of the process was paramount to how I analyzed and restoried participant truths. Like Jenoure (2002), parts of the restorying process felt “automatic” (p. 77).

Merriam and Tisdell (2016) also noted that research informed by postmodernism is “highly experimental, playful, and creative” (p. 11). Strictly speaking, all good research is creative, but this study certainly exploits and attempts to push the creative edge by employing

music as an agent of inquiry, analysis, and write-up. While I approached the topic of my study with seriousness, believing in its significance, I also subscribed to Kim's (2016) notion of aesthetic play. Kim defined the concept as "a play of ideas with an artistic meaning-making spirit that is playful and serious at the same time. It is an approach to narrative research design in which the narrative inquirer embodies curiosity and open-mindedness to the research process, allowing room for deliberation, intuition, and anticipation" (p. 301). For Kim (and me) this stance proved particularly useful during the data analysis phase of the research process. Like Kim, I "flirt(ed)" with the data (p. 187). I came with a beginner's mind. That is, I attempted to openly encounter the data with a mind as free as possible from my own presuppositions. I committed to a kind of light-hearted, tenacious curiosity that allows for surprise, uncertainty, and even disappointments (Kim, 2016). This stance, this researcher disposition, invites a type of full-bodied rigor that encompasses both head and heart (Leavy, 2015). It also serves as a form of bracketing.

Aesthetic Intersubjective Paradigm

While both constructivist and postmodern paradigms inform this research, a third, emerging paradigm perhaps most adequately describes the philosophical framework upon which this study is built. Understanding the need for paradigm building for this relatively new methodology, Chilton, Gerber, and Scotti (2015) articulated ontological, epistemological, and axiological presuppositions required for formal worldview building, which they named the aesthetic intersubjective paradigm. I subscribe to their assumptions and allowed them to guide my research.

Ontological pluralism. Like constructivist and postmodern paradigms, the aesthetic intersubjective paradigm is ontologically pluralistic; that is, in this paradigm, there are multiple

realities, truths. In this view, “people co-create or construct meaning — what they consider to be significant, real or true — through dialectical and dynamic processes” (p. 7). Truth is about the meanings that researchers and participants create in an intersubjective space, that is, in each other's presence.

Epistemological considerations. When one begins to consider the epistemological assumptions of the aesthetic intersubjective paradigm, its unique utility for arts-based researcher begins to emerge. It is the nature of *how* multiple realities can be known, aesthetically, that renders this paradigm particularly appropriate for arts-based research endeavors, like mine. Aesthetic knowings are preverbal, sensory, kinesthetic, and emotional. In this paradigm, research can communicate multiple truths in multiple ways. These pre-verbal, sensory kinesthetic, imaginal, spiritual and emotional ways of knowing “are often represented and expressed through dynamic and interactive symbols and metaphors” (p. 9). In subscribing to the idea that music can mediate meanings, I am calling upon this premise of aesthetic knowing.

Axiological considerations. The aesthetic intersubjective paradigm also addresses axiological issues, issues of value, by pointing to the healing and transformative power of the arts. According to Chilton, Gerber, and Scotti (2015), “This aesthetic intersubjective way of knowing — sensory, emotional, kinesthetic, imaginal, spiritual, preverbal — can lead to a more fulfilling, creative and free life in which awareness of self and others is expanded” (p. 12). In relation to my current arts-based study, the artistic renderings I created drew upon these ways of knowing and have the potential to raise critical consciousness and promote empathy each time they are performed. The research songs written through this study can enhance the quality of the lives of teachers who hear them. Teachers who have heard songs from the pilot study have

expressed that they related to them, that the process of listening to them was cathartic, and that the songs accurately reflected their experiences.

Position on the Art-Science Continuum

In addition to articulating paradigms that guide this study, I also want to situate this study along the art-science continuum (Leavy, 2017). As an arts-based researcher, I accept the notion that a strict divide between art and science is artificial. Barone and Eisner (2012) described the wall between art and science as porous; they postulated that within any well-wrought scientific process, artistic decisions are made. Similarly, in order to produce well-crafted art, artists engage in scientific tasks. Arts-based researchers can examine their philosophies and consider where their arts-based projects fall along that continuum. Projects located to the far left (arts side) of the continuum may utilize the arts to generate, analyze, and present data. Projects landing farther right (toward the science end of the continuum) incorporate the arts but as an addition to traditional qualitative methods. In these cases, the arts are used as a way to represent data that is garnered through traditional means (interview, observations, focus groups).

According to Leavy's (2017) characterization, this research falls roughly on the middle of the continuum because I almost equally embraced qualitative and arts-based practices. In the former camp, I used interviews, focus group, and narrative analysis in order to elicit and analyze data. In the latter, I gathered, analyzed, and represented data using musical coding, analysis, and sonic interpretation schema (Leavy, 2009). On paper, this study methodology is a near even split. However, when I consider this statement by Leavy, "Some practitioners adopt more of an arts-based practice and are likely to prioritize the insights gained from the acts of 'doing,' 'making' or 'experiencing' an arts practice" (p. 195), I place this study a bit farther to the left on the continuum. I *do* value the insights I gleaned and represented using musical inquiry. I also value

and validate the process by which they occurred and the ways in which the music and lyrics serve the research agenda.

Theoretical Framework

Cognitive Dissonance Theory

In addition to the philosophical stances outlined in the previous section of this document, this study was guided by a specific psychological theory: Festinger's (1957) Theory of Cognitive Dissonance. Festinger theorized that humans strive to maintain a peaceful mental state of homeostasis known as consistency. In order to perpetuate the preferred placid state, called consonance, people work to maintain unity between their beliefs, attitudes and actions (Festinger, 1957). If individuals become aware that they subscribe to opposing beliefs, a disconcerting cognitive conflict ensues. Similarly, consonance is disrupted if people are forced to act in ways that contradict their beliefs. To describe the resulting tension-filled state, Festinger coined the term cognitive dissonance.

Because people prefer consonance, and dissonance is uncomfortable, they attempt to navigate through the tension, that is, to return to consonance. Festinger posited three ways in which people resolve cognitive dissonance. Resolution can be achieved by changing a belief to match a behavior or a behavior to match a belief, gathering new data to potentially outweigh an existing belief, or reducing the importance of a belief. According to the theory, the desire to resolve dissonance and return to the harmonious state of constancy is a strong one — so strong that Festinger labelled it a drive.

Connection to Study

When teachers are asked to engage in pedagogical practices that contradict their beliefs about what constitutes sound pedagogy, they may experience cognitive dissonance. For example,

testing-related cognitive dissonance may arise if teachers are asked to narrow curricula or primarily engage in a single mode of curricular delivery. If teachers believe that a wide variety of subjects and pedagogical methods are called for, but they are asked or required to leave out certain subjects or utilize test-centric methods, they may experience a disconnect between their beliefs and their actions. Any ensuing discomfort constitutes cognitive dissonance. As the theory states, teachers in a dissonant state would necessarily need to resolve the tension or at least try to do so.

In the pilot I conducted, I studied three teachers who believed that standardized testing hampered their ability to teach in ways they believed were best for their students. Upon analyzing the data, I identified three potential expressions of cognitive dissonance. All three teachers claimed that the narrowed curricula they were asked to teach did not square with their desire to offer a robust curriculum. All shared concerns that their teaching efficacy faltered as a result of mandates to either teach to the test or, as one teacher put it, “teach with the test in mind.” Finally, all three questioned how to square their perceptions of their roles with administrators’ expectations to act mainly as agents of curricular delivery. I view these teachers’ concerns as examples of forced-compliance cognitive dissonance, which results from individuals being publically required to something they do not want to do.

Specific Methodological Applications

Cognitive dissonance theory meshes particularly well with the problem of practice in which this study is nested, namely that teachers who work in test-driven environments report a conflict between their belief about the practices they want to implement and the practices they are asked to enact (Abrams, Pedulla, & Madaus, 2003; Arroyo-Romano, 2016; Au, 2007; Berliner, 2011). Cognitive Dissonance Theory also aligns with this study’s stated purpose: To

create an in-depth description of K-12 teachers' pedagogical experiences in data-driven contexts. Given that pilot participants exhibited signs cognitive dissonance (i.e. agitation, discomfort) (Murray, 2007) and that these participants' experiences echoed those reported in the literature, using this theory to guide the research is a logical choice.

I used Cognitive Dissonance Theory to mediate several of this study's methods. First, the theory informed my researcher-created interview and focus group protocols. When writing these protocols, I composed some questions to search for potential evidence of cognitive dissonance. I created other questions to elicit responses that specifically reference internally-inconsistent beliefs/attitudes or belief/action disparities. A final set of protocol questions were aimed at discovering the methods participants might be using to navigate any curricular-related cognitive dissonance.

During data analysis, Cognitive Dissonance Theory acted as an interpretative lens. As I coded transcript data, I searched for evidence of cognitive dissonance. While coding, I noted specific examples of dissonant beliefs, attitudes, and belief/action pairs. I coded for key ideas that might be verbal signposts of cognitive dissonance (frustration, worry, doubt, concern). Finally, I sifted the data for evidence of ways that teachers attempt to navigate testing-related cognitive dissonance, whether it be by changing beliefs/behaviors, gathering new data to potentially outweigh an existing belief, or reducing the importance of a held belief.

Finally, this framework guided the data write-up portion of this research. When I discovered evidence of cognitive dissonance, I included that finding in either the program notes or songs that I wrote. Because this arts-based research employs musical inquiry, it opens up an intriguing possibility to novelly extend Festinger's theory. Festinger chose to frame homeostasis-seeking in terms of consonance and dissonance. Whether knowingly or not, when Festinger

substituted the word “consonance” for constancy and coined the phrase “cognitive dissonance,” he employed well-known musical terminology. In Western music theory, when two or more pitches are consonant, the ear codes these harmonies as pleasant (LoPresto, 2015). Conversely, dissonant harmonies can jar and unsettle the listener. Depending upon desired emotional effects, composers employ both consonant and dissonant harmonies. In Western, tonal music, dissonant harmonies are used but often as harmonies to move *through* . . . to resolve.

Since I wrote narrative data songs, I exploited the connections between Festinger’s terms and their musical counterparts. When a participant demonstrated curricular-related cognitive dissonance, I rendered that state both harmonically and lyrically. The result is a musical narrative that I intend to literally ring true to the participants’ experiences. This novel application of the theory remains true to the theoretical framework while simultaneously exploiting its metaphorical implications. Moreover, because I curated the musical narratives via a series of reflective program notes, I was able to provide some examples of how I wrote dissonance into the songs.

Research Practices

Participants and Recruitment

Using purposeful sampling, I identified and recruited four local K-12 educators to participate in this study. I recruited four participants in order to gain insights from multiple perspectives. This decision allowed for the possibility that, should a teacher have chosen to discontinue study participation (none did), I would still have at least three participants and thus, multiple perspectives upon which to draw. In addition to working locally, teachers eligible to participate in this study must have been practicing for at least one year, been practicing during

the data collection phase of this study, and be working in a public or charter school in the greater New Orleans area. All four teachers met these inclusion criteria.

I recruited for this study in several ways. First, I asked gatekeepers to connect me with currently practicing teachers either enrolled in or recently graduated from local teacher training programs. I also recruited from the community at large through advertising this study online. In conjunction with purposeful sampling, I recruited participants through both convenience and snowball sampling. Since I have taught courses at the local university and maintain connections with teachers whom I have taught, I approached these educators and asked if either they or eligible teachers they know would be willing to participate in this study.

Data Collection

Semi-structured interviews. In an effort to address the research agenda, I conducted one semi-structured, in-depth interview with each of the study participants. The in-depth interview took place prior to end-of-year, standardized testing. According to BESE (“2017-2018 Assessment Calendar,” 2017), K-12 end-of-year testing took place between April 9th and May 18th of 2018 (depending upon the grades and type of test). Schools published their own testing calendars, so I coordinated interview dates with participants based upon their schools’ particular schedule. All in-depth interviews took place prior to the beginning of the testing window. The focus group took place while some participants were in the midst of testing; for others, the window had just closed.

Using the research questions and the guiding theory, I composed interview protocol questions. While some questions asked to participants were identical, others were context-specific (e.g. depending upon grades and subjects taught). Because these interviews were semi-structured, the questions I asked were open-ended, and I formulated follow-up, probing questions

in real time meant to elicit deeper understandings of the phenomena under study (Merriam & Tisdell, 2016). All questions were meant to “yield descriptive data, even stories about the phenomenon” (Merriam & Tisdell, 2016, p. 120). During the interviews, which were recorded and later transcribed, I took research notes in order to better frame follow up questions.

Focus group. After all in-depth interviews were completed, I conducted a focus group with three of the four study participants. Like the interviews, the focus group was semi-structured. I framed focus group questions, but I also engaged participants in conversations with each other. The participants dialogued extensively and posed questions to each other during the course of the focus group. In addition to posing follow up questions based upon interim data analysis of the semi-structured interview, I also utilized an arts-based data gathering technique. I performed two songs that were written as the results of the pilot study and asked participants to respond to the songs by underlining and discussing portions of the songs that they related to or wanted to discuss. Finally, each focus group participant participated in a writing exercise meant to encourage reflection around the ways that they practice in light of standardized testing. Each participant read their writing aloud and the participants responded to each others’ writing. Data gathered from this recorded focus group were transcribed, analyzed, and incorporated into the final write up.

Email follow up. On several occasions, in order to obtain follow up data, I contacted participants via email. Email was a practical way to obtain follow up data as participants were able to answer questions at times that were convenient to them. In addition to my eliciting data via email, on two occasions, participants emailed me unsolicited follow up thoughts. I included these thoughts as study data and analyzed them according to the previously outlined protocol.

Data Analysis

Narrative analysis. According to Kim (2016), narrative research seeks to restory data as a cohesive work that “integrates events and happenings into a temporally organized whole with a thematic thread” (p. 197). To that end, I analyzed the data using Polkinghorne’s (1995) narrative mode of analysis. Such analysis requires that data be analyzed recursively, moving between parts and whole, external events and participant reactions to those events.

Problem-solution coding. Narrative research places participants in the role of protagonists (Kim, 2016). As such, the potential existed that the participant-protagonists would encounter difficulties to overcome (i.e. problems to solve). Ollerenshaw and Creswell (2002) described a problem-solution mode of narrative analysis, which encourages researchers to uncover problems a participant may face. Since cognitive dissonance is definitionally a problem to be solved, problem-solution coding was a logical fit for this study. To code in this way, I delineated sections of the data as problem-specific and solution-specific. Within these sections, I coded for specific problems as well as for ways the participants attempted to solve any curricular-related problems (including, but not limited to, cognitive dissonance.)

Arts-based coding. For the arts-based components of this research, I used songwriting as a mode of analysis and data rendering; I coded data using a modified form of performance collage. According to Leavy (2015), performance collage involves “musically coding and writing up data culminating in a musical performance” (p. 133). Specifically, in order to musically restory the data, the data indicative of significant findings were coded for artistic/musical potential. I coded data for words/phrases that are rich in sensory detail or are rife with lyrical potential (e.g. image-laden text, metaphor, and simile). Doing so aligns with the

propensity of narrative researchers to incorporate literary devices in their writings (Clandinin & Connelly, 2000).

Songwriting. Songs were the final product of this research. Songwriting also acted as a mode of data analysis. In the process of writing songs, I set participants experiences particular musical ways (i.e. refrains, bridges, ascending melodies, dissonant/consonant harmonies). Each of these musical decisions required me to constantly revisit the data, reconsider its plot, themes, and, at times, my previous interpretations. In this way, the process of art-making deepened the data analysis. Arts-based-researcher and pedagogue Peter Gouzouasis posited that arts-based researchers create a hermeneutic circle between their artistic process and their research agenda (P. Gouzouasis, personal communication, November 4, 2016). He posed two questions: “How does the research serve the art and how does the art serve the research?” In my data analysis, I weighted the second question heavier, striving to ensure that the art serves the research agenda well.

Data Representation

While the songwriting process acted as a mode of data analysis, a cycle of narrative songs was the primary form of data representation (i.e. the findings) of this study. All musical renderings of the data restoried the data in two ways: verbally (lyrically) and sonically. As a trained musician, I have at my disposal knowledge of musical tools: Melody, harmony, rhythm, form, dynamics, and arrangement strategies. I used this knowledge to render songs that invite listeners to enter into the participants’ experiences.

This type of musical data portrayal, what Leavy (2015) referred to as sonic interpretation schema has, to my knowledge, never been utilized to conduct doctoral level research. I have unearthed three dissertations (Bakan, 2014; Carson, 2017; Viega, 2013) that incorporate music to

varying degrees. Carson's hip-hop dissertation is the closest in form (completely arts-based) to the research I conducted but to date, I have yet to discover researchers undertaking songwriting in the ways this research describes. Carson wrote a 34 track hip-hop album, but his study included no participants. Bakan wrote a single autoethnographic song, which he included in his dissertation. Viega digitized and performed the hip-hop pieces composed by his music therapy clients. However, using participant data to dictate the songs I write is, to my knowledge, a novel approach in my discipline.

As previously stated, the data informed the form each song took. However, as a pianist and singer trained in classical and popular traditions, these influences certainly came into play. They, along with my singing voice and particular pianistic style, imbued the songs with my particular researcher thumbprint (Barone & Eisner, 2012; Leavy, 2015).

Nearly twenty years ago, Clandinin and Connelly (2000) invoked narrative researchers to "work at the boundaries, to stretch themselves in new ways as they try to come close to understanding experience" (p. 189). These narrative songs, written from the perspectives of the teachers themselves, tap into the "imaginative possibility" (p. 189) to which Clandinin and Connelly referred. Writing up findings as songs allowed me to access aesthetic knowledge and convey participants' truths in unique, evocative ways, ways that Barone and Eisner (2012) claimed "cannot be secured in non-music forms" (p. 1).

By bringing research consumers close, these ways offer listeners "possibilities for reliving, for new directions" (Clandinin & Connelly, 2000, p. 189). Much of my data write-up was music. The music I wrote was scholarship enacted. However, I recognized a need to answer the question Eisner (2008) raised, "Will the images made through arts-based research possess a sufficient degree of referential clarity to engender a common understanding of the situation being

addressed?” (pp. 18-19). Since I did not conduct art for art’s sake but in order to analyze data and disseminate findings, I needed to make sure to provide enough referential clarity for the research audience to understand the findings I composed. Thus, I curated the musical results of this research via a set of program notes. The notes included necessary background information about each teacher who participated in this study. In the program notes, I fleshed out the context out of which each song arose, the problem or solution each song addressed, as well as provided a brief musical analysis of each song.

Evaluation

In regards to evaluating applied research, Merriam and Tisdell (2016) wrote the following: “To have any effect on either the practice or theory of a field, research studies . . . need to present insights and conclusions that ring true to readers, practitioners, and other researchers” (p. 238). In the case of this research, those words, “ring true” are literal. Those who listened to this research with an ear to evaluate needed be able to do so with criteria that are attuned to its methodology. Evaluative terms (reliability, validity, generalizability) originally linked to quantitative methodologies have been redefined to fit qualitative forms of research. However, Wolcott (1994) argued that these terms do not apply to qualitative research in general. Creswell (2013) and Connelly and Clandinin (1990) stated that narrative studies should be evaluated differently than other forms of qualitative research (e.g. on how engaging the story is). Even in qualitative circles, criteria other than reliability, validity, and generalizability have been posited. As arts-based research has become more widely utilized, arts-based scholars have proposed methodological-specific, evaluative criteria (Barone & Eisner, 2012; Leavy, 2015) I proposed using Barone and Eisner’s (2012) as a guide for evaluating this arts-based dissertation. Table 1 summarizes their criteria.

Barone and Eisner's (2012) Evaluative Criteria (pp. 148-153)
Incisiveness: The work cuts to the core of the research issue.
Concision: The work occupies the minimal amount of space/uses the least amount of verbiage necessary to serve its purpose.
Coherence: The features of the work hang together as a strong form.
Social Significance: The work focuses on issues that make a sizable difference in the lives of people within a society.
Generativity: (not to be confused with generalizability) The work raises larger questions, extends beyond the single instance being studied. From the particular instance, questions/observations about universals are raised.
Evocation and Illumination: The work evokes feeling in the consumer and sheds light on an object or process by helping readers see them in new ways.

Table 1. Evaluative Criteria for Arts-Based Research

Researcher Positionality

A/r/tographer

I approached this study from a stance that is particular to arts-based educational researchers, the position of an a/r/tographer. A/r/tography is a metaphor for the artist-teacher-researcher identity that some art-based educational researchers occupy (Leavy, 2015; Sinner, Leggo, Irwin, Gouzouasis, & Grauer, 2006; Springgay, Irwin & Kind, 2005; Springgay, Irwin, Leggo, & Gouzouasis, 2008). It is a metaphor that fits all aspects of my researcher identity. Before I was a researcher, I was a teacher. Before I was a teacher, I was a musician. Rather than abandon these aspects of my identity, I embraced them; I brought them to my research with the belief that far from hindering, they would aid me through all phases of the research process. As

an artist, I have been trained to make meaning in musical ways. As a teacher, I am privy to a certain body of insider discourse that helps me select and frame pertinent issues. As a fledgling researcher, I am acquiring research dispositions and skills that help me systematically inquire and report the results of my inquiry. Arts-based research allowed me to bring these three aspects of my identity to bear upon this research in order to create novel, engaging research products that address significant research issues.

Insider-Outsider

Dwyer and Buckle (2009) defined insider researchers as members of the community they study. Because I am a trained educator with classroom experience, I am a member of the teaching profession. As such, I'm privy to the discourse-specific knowledge of the teaching community. I can "talk the teacher talk." However, I am not currently teaching in a K-12 setting. Moreover, I won't be working at any of the schools where my participants practice, so in this way I exist on the outside of their particular community. For these reasons, I position myself as an insider-outsider.

Ethical Considerations

As I embarked upon this research, I had several ethical considerations to address. First, prior to commencing with this study, I obtained approval from the University of New Orleans institutional Review Board. After obtaining IRB approval, I obtained the written informed consent of all study participants. During that process, I fully explained this study's purpose and protocol and let my participants know that they may discontinue participation in this study at any time with no negative consequences. All participants chose to participate for the duration of the study and to allow me to publicly perform the results of this study. However, one participant was unable to attend the focus group due to a scheduling conflict.

In addition to obtaining informed consent and IRB approval, I took the standard precautions to protect my participants' identities. I did so by using password-protected documents for the data I collected. When I wrote up the data, I used participant-approved pseudonyms for my study participants and referred to their teaching contexts in general terms such as "a local charter school" instead of by name.

While I am not working directly with a vulnerable population, the teachers who participated in this study do have contact with minors in the K-12 schools where they work. In the pilot of this study, teachers in a focus group mentioned students by name, and the stories they told about specific students yielded themes that I believed might arise again in this study. As I suspected, they did, so in order to protect those students' identities, I did not refer to any students by name. I also vetted those students' stories with the study participants in order to make sure that no identifying information would be present in the final write up.

A hallmark of qualitative research is thick description, and my goal with this narrative was to conduct a nuanced deep dive into the lived experiences of the teachers who operate in schools/systems where they are being asked to engage in test-related practices they may find distasteful. Inherent in this exploration was the potential unmasking of systemic power dynamics. Teachers did discuss how those dynamics played out in their practice, and I wrote about them in several songs. I can see how these dynamics could have been cause for concern to my study participants. Teachers I previously interviewed indicated that they were fearful of disrupting the system by delivering broad curricula in their classrooms. The potential existed for my study to create an issue for teachers operating in what they deem to be a hostile teaching environment. Teachers who believe that they must acquiesce or potentially be reprimanded by (or even fired) by their school leaders for failing to conform to the mandated curricula could

have been particularly afraid of their identities being revealed. I saw a tension between richly rendering the experiences of these teachers and protecting their identities. Obviously, the ethical considerations trumped the research agenda. At times, I needed to scale back the stories I rendered and leave out specific details in order to protect the privacy of my participants. Since I did address power dynamics in several songs, I made sure to vet these songs with teachers prior to performing them live.

The privacy issues were particularly at the forefront of my ethical considerations because this dissertation is performative. I disseminated the study results via a live musical performance of the data songs I wrote. The audience of this research was not just my dissertation committee; it was anyone in the city who attended the performance. Since I advertised and garnered an audience, the potential for someone in power to see the results was more likely than with a traditional dissertation.

In order to address this ethical concern, I informed my participants that local stakeholders might see or hear about the performance of this dissertation. I also member checked my findings individually with each participant. I did so not only to vet for accuracy but also to vet for potential identifying information present in the research products. After composing drafts of the narrative songs, I met individually with each participants in order to member check. Two participants asked that I make minor lyrical changes (which I did). Otherwise, the participants assured me that the songs accurately represented their viewpoints. In addition to vetting for accuracy, I also asked participants if I had left anything out of their narrative that they wanted me to include. Each participant agreed that the narratives I constructed did not leave out any information they deemed important. Finally, I gave each participant the opportunity to vet the program notes for accuracy.

Limitations

This study is an in depth, context-specific examination of four K-12 teachers' experiences with standardized test-related pedagogical phenomenon. Any findings related to this study are not generalizable. However, Barone and Eisner (2012) moved the evaluative discussion of arts-based research away from issues of generalizability and asked instead whether a study is generative. The question is not whether a study's results can be replicated. Rather, it is, "Does the research raise larger questions that extend beyond the single issue studied?" (See table 1).

This study was both limited and enhanced by my background as a teacher and doctoral student in curriculum and instruction. My practice and study have helped me construct study-pertinent schema that aided my in-depth analyses of the data and pertinent literature. However, those same schema have also enabled me to construct biases about standardized testing and the types of curricula that I believe should be taught. I do believe that a broad curriculum serves students best. I do think that situated, differentiated, culturally-relevant pedagogies are core tenets of sound pedagogy. I attempted to bracket these biases by observing them and by writing about them in a researcher journal and in the program notes. Keeping in mind that my research is meant to be an accurate reflection of the participants' experiences, I recognized that my bias could have bled into the data interpretation. Therefore, the in-depth, in-person member checking in which I engaged with each participant, was a particularly important part of bias mitigation.

References

- “About charter schools” (2017). Louisiana Association of Public Charter Schools. Retrieved from <https://lacharterschools.org/about-charter-schools/>
- “Accountability, school performance scores, combination schools” (2015). Retrieved from <https://www.louisianabelieves.com/docs/default-source/accountability/summit-presentati on-sps--combination-schools.pdf?sfvrsn=2>
- Abrams, L. M., Pedulla, J. J., & Madaus, G. F. (2003). Views from the Classroom: Teachers’ opinions of statewide testing programs. *Theory into Practice*, 42(1), 18-29.
- Agee, J. (2004). Negotiating a teaching identity: An African American teacher's struggle to teach in test-driven contexts. *Teachers College Record*, 106(4), 747-774.
- Alford, C. (2010). *Cognitive dissonance experienced by general education teachers when teaching in inclusion classes*. (Unpublished doctoral dissertation). University of Phoenix. Phoenix, AZ.
- Alocer, P. (2017). History of standardized testing in the United States. *National Education Association*. Retrieved from <http://www.nea.org/home/66139.htm>
- Amrein, A. L., & Berliner, D. C. (2002). *An analysis of some unintended and negative consequences of high-stakes testing*. Education Policy Research Unit, Education Policy Studies Laboratory, College of Education, Division of Educational Leadership and Policy Studies, Arizona State University.
- Amrein, A.L., & Berliner, D. C. (2003). The Effects of high-stakes testing on student motivation and learning. *Educational Leadership*, 60(5), 32-38.
- Anagnostopolous, D. (2003a). The new accountability, student failure, and teachers' work in urban high schools. *Educational Policy*, 17(3), 291-316.

- Anagnostopolous, D. (2003b). Testing and student engagement with literature in urban classrooms: A multi-layered perspective. *Research in the Teaching of English*, 38(2), 177-212.
- Arroyo-Romano, J. (2016). The ethical teacher in a high-stakes testing environment. *Journal of Studies in Education*, 6(2), 56-74.
- Au, W. (2007). High-stakes testing and curricular control: A qualitative metasynthesis. *Educational Researcher*, 36(5), 258-267.
- Au, W. (2013). Hiding Behind High-Stakes Testing: Meritocracy, objectivity, and inequality in U.S. education. *The International Education Journal: Comparative Perspectives*, 12(2), 7-19.
- Augustine, E. (2016) *Teachers' perspectives about classroom dynamics shaped by standardized testing in k-12 classrooms of Indiana: a narrative study* (Unpublished doctoral dissertation). Ball State University, Muncie, IN.
- Bakan, D. (2014). *A song of songs: A/r/tography, autoethnography, and songwriting as music education research*. (Unpublished doctoral dissertation). University of British Columbia, Canada.
- Barone, T., & Eisner, E. (1997). Arts-Based educational research In. M. Jaeger (ed.) *Complementary Methods in Research Education*, 2nd edition. 73-116. Washington, DC: American Educational Research Association.
- Barone, T., & Eisner, E. W. (2012). *Arts based research*. Los Angeles, CA: SAGE.
- Barton, K. C. (2005). "I'm not saying these are going to be easy": Wise practice in an urban elementary school. In E. A. Yeager & O. L. Davis Jr. (Eds.), *Wise social studies teaching in an age of high-stakes testing* (p. 11-31). Greenwich, CT: Information Age

Publishing.

- Bell-Weixler, L., Harris, D., Barrett, N. (2017). Teachers' perspectives on the learning and work environments under the New Orleans school reforms. *Education Research Alliance for New Orleans*, Retrieved from <https://educationresearchalliancenola.org/files/publications/Bell-Weixler-Harris-Barrett-Teachers-Perspectives-on-the-Learning-and-Work-Environments-Under-the-New-Orleans-School-Reforms.pdf>
- Berg, M. (2006). *The costs of success on the high-stakes standardized reading test in terms of standards based instruction in a high-minority, high poverty elementary school.*
- Berliner, D. (2011). Rational responses to high stakes testing: the case of curriculum narrowing and the harm that follows. *Cambridge Journal of Education*, 41(3). 287-302.
- Blazer, S. (2011). Unintended consequences of high-stakes testing. *Information Capsule: Research Services*, 1008, 1-12.
- Bolgatz, J. (2006). Revolutionary talk: Elementary teacher and students discuss race in a social studies class. *Social Studies*, 96, 259–264.
- Bordelon, S. J. (2010). Making the grade? A report card on special education, New Orleans charter schools, and the Louisiana charter schools law. *Loyola Journal of Public Interest Law*, 11(2), 441-449.
- Braun, H., Chapman, L. & Vezzu, S. (2010). The black-white achievement gap revisited. *Education Policy Analysis Archives*, 18(21), 1-99.
- Brinson, D., Boast, L., Hassel, B. C., & Kingsland, N. (2012). New Orleans–style education reform: A guide for cities: Lessons learned, 2004–2010. New Orleans, LA: New Schools for New Orleans.

- Brimijoin, K. (2005). Differentiation and high-stakes testing: An oxymoron? *Theory Into Practice*, 44(3), 254-261.
- Buras, K. L. (2014). Charter schools, race, and urban space: Where the market meets grassroots resistance. New York, NY: Routledge.
- Campbell, E. (2003). *The ethical teacher*. Philadelphia, PA: McGraw-Hill Education.
- Carson, A.D. (2017). *Owning my masters: The rhetorics of rhymes & revolutions*. (Unpublished doctoral dissertation). Clemson University, Clemson, SC.
- Chait, J. (2015, August). How New Orleans proved urban-education reform can work. *New York Magazine*. Retrieved from <http://nymag.com/daily/intelligencer/2015/08/how-new-orleans-proved-education-reform-can-work.html>
- Chilton, G., Gerber, N., & Scotti, V. (2015). Towards an aesthetic intersubjective paradigm for arts-based research: An art therapy perspective. *UNESCO Observatory Multi-disciplinary Journal in the Arts*, 5(1), 1-27.
- Chudowsky, N., & Chudowsky, V. (2009). State test score trends through 2007-8, part 2: Is there a plateau effect in test scores? *Center on Education Policy*. Washington, DC.
- Ciolino, M., Kirylo, J., Mirón, L., & Frazier, K. (2014). Education reform in New Orleans: voices from the Recovery School District. *Policy Futures in Education*, 12(4). 463-467.
- Clandinin, F., & Connelly, F. (2000). *Narrative inquiry: Experience and story in narrative Research*. San Francisco: Jossey-Bass.
- “Comments and questions to the draft ESSA state plan prepared by the Louisiana Department of Education” (2017). Retrieved from <https://www.louisianabelieves.com/docs/default-source/louisiana-believes/feedback-from-education's-next-horizon.pdf?sfvrsn=3>

- Connelly, F., & Clandinin, D. (1990). Stories of Experience and Narrative Inquiry. *Educational Researcher*, 19(5), 2-14.
- Costigan, A. (2008). Canaries in the coal mine: Urban rookies learning to teach language arts in “high priority” schools. Canaries in the Coal Mine: Urban Rookies Learning To Teach Language Arts in “High Priority.” *Teacher Education Quarterly*, 35(2), 85-103.
- Craig, C. J. (2004). The Dragon in School Backyards: The Influence of mandated testing on school contexts and educators’ narrative knowing. *Teachers College Record*, 106(6), 1229-1257.
- Creswell, J. (2013). *Qualitative inquiry & research design* (3rd ed.). Thousand Oaks, CA: Sage.
- Croco, M., & Costigan, A. (2007). The narrowing of curriculum and pedagogy in the age of accountability: Urban educators speak out. *Urban Education*, 42(6), 512-535.
- Csikszentmihalyi, M. (1990). *Flow*. New York, NY: Harper Collins.
- Decuir, E. (2012). *Louisiana Educational Assessment Program (LEAP): A historical analysis of Louisiana's high stakes testing policy*. (Unpublished doctoral dissertation). Georgia State University, Athens, GA.
- Dixson, A., Buras, K., & Jeffers, E. (2015). The color of reform: Race, education, and charter schools in post-Katrina New Orleans. *Qualitative Inquiry*, 21(3), 288-299.
- Dreilinger, D. (2016, August, 2016) “Louisiana's younger students make progress on 2016 tests.” *Nola.com*. Retrieved from http://www.nola.com/education/index.ssf/2016/08/Louisiana_test_scores_2016.html
- Duffy, G. (2007). Thriving in a high-stakes testing environment. *Journal of Curriculum and Instruction*, 1(1), 7-13.

- Dwyer, S., & Buckle, J. (2009). The space between: On being an insider-outsider in qualitative Research. *International Journal of Qualitative Methods*, 8(1), 54-63.
- Eisner, E. (2008) Persistent tensions in arts-based research. In Cahnmann-Taylor, M., Siegesmund, R. (Eds). *Arts-based research in education*. (pp. 16-27). New York, NY: Routledge.
- Elementary and secondary education act of 1965*. P.L. 89-10.
- Every student succeeds Act of 2015*. P.L. 114-95.
- Ferguson, M., Kober, N. & Rentner, D. (2016). Listen to us: Teacher views and voices. *Center On Education Policy*. Retrieved from <https://files.eric.ed.gov/fulltext/ED568172.pdf>
- Festinger, L. (1957). *A theory of cognitive dissonance*. Evanston, IL: Row & Peterson.
- Freire, P. (2001). *Pedagogy of freedom: Ethics, democracy, and civic courage*. Lanham, MD: Rowman & Littlefield.
- Freire, P. (2004). *Pedagogy of the oppressed, 30th anniversary edition*. New York, NY: The Continuum International Publishing Group.
- Garda, R. (2010). The politics of education reform: Lessons from New Orleans. *Journal of Law & Education*, 40(1), 1-42.
- Gerwin, D., & Visone, F. (2006). The freedom to teach: Contrasting history teaching in elective and state-tested courses. *Social Education*, 34(2), 259-282.
- “Grades 3-8 assessments” (2017). Department of education: Louisiana believes. Retrieved from <http://www.louisianabelieves.com/measuringresults/grades-3-8-assessments>
- Graham, R. (2007). Theory of cognitive dissonance as it pertains to morality. *Journal of Scientific Psychology*, June 2007, 20-23.

- Gouzouasis, P. (2016, November 4), personal communication.
- Guest, G., Namey, E. & Mitchell, M. (2012). Qualitative research: defining and designing. In *Collecting qualitative data: a field manual for applied research*. Los Angeles, CA: SAGE.
- “High school assessments” (2017). Retrieved from <http://www.louisianabelieves.com/measuringresults/assessments-for-high-schools>
- Hillocks, G., Jr. (2002). *The testing trap: How state writing assessments control learning*. New York: Teachers College Press.
- Hinde, E. (2003). The tyranny of the test: Elementary teachers’ conceptualizations of the effects of state standards and mandated test on their practice. *Current Issues in Education*, 6(10).
- Holley-Walker, D. (2007). The accountability cycle: the Recovery school district act and New Orleans’ charter schools, *Connecticut Law Review*, 40(1), 125-163.
- Hull, A.M. (2002). Neuroimaging findings in post-traumatic stress disorder. *The British Journal of Psychiatry*, (181) 102-10.
- Improving America’s schools act of 1994*. P.L.103-382.
- Jerald C. D. (2006). The hidden costs of curriculum narrowing. *The Center for Comprehensive School Reform and Improvement*. Washington, DC. Retrieved from <http://www.centerforcsri.org/files/CenterIssueBriefAug06.pdf>
- Jenoure, T. (2002). Sweeping the temple. In C. Bagley & M. Cancienne (Eds.) *Dancing the Data*. New York, NY: Peter Lang Publishing.
- Jewson, M. (2017). Compare 2017 school performance scores for New Orleans schools. *The Lens*, November 7th. Retrieved from <https://thelensnola.org/2017/11/07/see-the-2017-school-performance-scores-all-in-one-place/>

- Jones, B., & Egley, R. (2004). Voices from the frontlines: Teachers' perceptions of high-stakes testing. *Education analysis policy archives*. 12(39), 1-34.
- Jones, M., Jones, B., Hardin, D., Yarbrough, T., & Davis, M. (1999). The impact of High-stakes testing on teachers and students in North Carolina. *Phi Delta Kappa International*, 81(3), 199-203.
- Kim, J. (2016). *Understanding narrative inquiry: the crafting and analysis of stories as research*. Los Angeles: SAGE.
- King K. V., & Zucker S. (2005). *Curriculum narrowing*. San Antonio, TX: Harcourt Assessment.
- Knowles, J.G., & Cole, A.L. (2008). *Handbook of arts in qualitative research*. Thousand Oaks, CA: Sage Publications.
- Ladson-Billings, G. (1995). But that's just good teaching! The case for culturally relevant pedagogy. *Theory Into Practice*, 34(3), 159–165.
- Ladson-Billings, G. (2009). *The dreamkeepers: Successful teachers of African American children*. San Francisco, CA: Jossey-Bass.
- Leavy, P. (2013). *Fiction as research practice*. New York, NY: Taylor & Francis.
- Leavy, P. (2015). *Method meets art arts-based research practice*. New York, NY: The Guilford Press.
- Leavy, P. (2017). *Handbook of research design: Quantitative, qualitative, mixed-methods, arts-based, and community-based participatory research practices*. New York, NY: Guilford Press.

- Libresco, A. S. (2005). How she stopped worrying and learned to love the test. . . sort of. In E. A. Yeager & O. L. Davis Jr. (Eds.), *Wise social studies teaching in an age of high-stakes testing* (pp. 33-49). Greenwich, CT: Information Age Publishing.
- Lipman, P. (2011). *The new political economy of urban education: Neoliberalism, race, and the right to the city*. New York, NY: Routledge.
- LoPresto, M. (2015). Measuring musical consonance and dissonance. *The Physics Teacher*, 53(225), 55-60.
- Madaus, G., & Russell, M. (2011). Paradoxes of high-stakes testing: High-stakes testing and instruction. *Peabody Journal of Education*, 190(1/2), 21-30.
- Merriam, S., & Tisdell, E. (2016). *Qualitative research: A guide to design and implementation*. San Francisco, CA: Jossey-Bass.
- McMurrer, J. (2007). NCLB year 5: Choices, changes, and challenges: curriculum and instruction in the NCLB era. U.S.A.: Center on Educational Policy.
- National Board on Educational Testing and Public Policy. (2003). *Perceived effects of state-mandated testing programs on teaching and learning: Findings from a national survey of teachers*. Chestnut Hill, MA: Boston College.
- National defense of education act of 1958*, P.L. 85-864.
- “National defense of education act,” (2017). History, Art & Archives: United States House of Representatives. Retrieved from <http://history.house.gov/HouseRecord/Detail/15032436195>
- Newberg-Long, D.K. (2010). *Narrowing of curriculum: Teaching in an age of accountability*. (Unpublished doctoral dissertation). University of Denver, Denver, CO.

Nichols, S.L., & Berliner, D.C. (2008). Testing the joy out of learning. *Educational Leadership*, 65(6), 14-18.

Nobles III, W. (2017). Louisiana schools' 2017 performance score rises 3.8 points to B grade.

The Times Picayune, November 7th, Retrieved from. http://www.nola.com/education/index.ssf/2017/11/2017_school_performance_scores.html

No Child Left Behind Act of 2002, P.L. 107-110, 20 U.S.C. § 6319.

Noddings, N. (1984). *Caring: A feminine approach to ethics and moral education*. Berkeley, University of California Press.

Ollerenshaw, J. A., & Creswell, J. W. (2002). Narrative research: A comparison of two restorying data analysis approaches. *Qualitative Inquiry*, 8(3), 329-347.

Perkins, G., & Wellman, N. (2008). The impact of high-stakes testing: Recommendations for professional school counselors and other support personnel. *Academic Leadership*, 8(3) 1-6.

Phelps, R. (2017). *Kill the Messenger: The war on standardized testing*. New York, NY: Routledge.

Polkinghorne, D. (1991). Narrative and self concept. *Journal of Narrative and Life History*. 1(2 & 3), 135-153.

Polkinghorne, D. (1995). Narrative configuration in qualitative analysis. *International Journal of Qualitative Studies in Education*, 8(1), 5-23.

Primary sources: America's teachers on America's schools. (2010). U.S.A.: Bill & Melinda Gates Foundation/Scholastic.

- Primary sources: America's teachers on the teaching profession. (2012). U.S.A.: Bill & Melinda Gates Foundation/Scholastic.
- Public agenda reality check 2000. *Education Week*, XVIII (43), 1-8.
- Rapp, D. (2002). National Board-certified teachers in ohio give state education policy, classroom policy, and high-stake testing a grade of f. *Phi Delta Kappan*, (84)3, 215.
- Reese, M., Gordon, S., & Price, L. (2004). Teachers' perceptions of high stakes testing. *Journal of School Leadership*. 14(5). 464-496.
- Rentner, D. S., Scott, C., Kober, N., Chudowsky, N., Joftus, S., & Zabala, D. (2006). *From the capital to the classroom: Year 4 of the no child left behind act*. Washington, DC: Center on Education Policy.
- Richter, D. (2017a). Now that the testing is over. Unpublished manuscript.
- Richter, D. (2017b). Arts-based research emerging: A generals portfolio. Unpublished manuscript.
- Rolling, J. H. (2013). *Arts-based research primer*. New York: Peter Lang Publishing.
- Santelices, M., & Wilson, M. (2010). Unfair treatment? The case of Freedle, the SAT, and the standardization approach to differential item functioning. *Harvard Educational Review*, 80(1), 106-133.
- Santoro, D. (2011). Good teaching in difficult times: demoralization in the pursuit of good work. *American Journal of Education*, 118(1), 1-23.
- Santoro, D. (2011a). Teaching's conscientious objectors: Principled leavers of high-poverty schools. *Teachers College Record*, 113(12), 2670-2704.
- Segall, A. (2003). Teachers' perceptions of the impact of state-mandated standardized testing: The Michigan Educational Assessment Program (MEAP) as a case study of

- consequences. *Theory and Research in Social Education*, 31(3), 287-325.
- Sinner, A., Leggo, C., Irwin, R. L., Gouzouasis, P., & Grauer, K. (2006). Arts-based educational research dissertations: Reviewing the practices of new scholars. *Canadian Journal of Education/Revue Canadienne De L'éducation*, 29(4), 1223.
doi:10.2307/20054216
- Smagorinsky, P., Lakly, A., & Johnson, T. S. (2002). Acquiescence, accommodation, and resistance in learning to teach within a prescribed curriculum. *English Education*, 34(3), 187-213.
- Smith, A. M. (2003). Negotiating control and protecting the private: History teachers and the Virginia standards of learning. In S. G. Grant (Ed.), *Measuring history: Cases of state-level testing across the United States* (pp. 221-247). Greenwich, CT: Information Age Publishing.
- Sondel, B. (2016) "No Excuses" in New Orleans: The silent passivity of neoliberal schooling. *The Educational Forum*, 80(2) 171-188.
- Springgay, S., Irwin, L., & Kind, S. (2005). "A/r/tography as living inquiry through art and text." *Qualitative Inquiry* 11(6), 897-912.
- Springgay, S., Leggo, C., & Gouzouasis, P. (2008) (Eds.) *Being with a/r/tography*, Rotterdam: Sense Publishers.
- Strunk, O., Barrett, N., & Lincove, J. (2017). When tenure ends: Teacher turnover in response to policy changes in Louisiana. *Policy Brief: Education Research Alliance for New Orleans*. New Orleans, LA. Retrieved from <https://educationresearchalliancencola.org/files/publications/022217-Strunk-Barrett-Arnold-Lincove-When-Tenure-Ends-Teacher-Turnover-in-Response-to-Policy-Changes-in-Louisiana.pdf>

- “Student enrollment and demographics” (2017). Department of Education. Retrieved from <https://www.louisianabelieves.com/docs/default-source/katrina/final-louisiana-believes-v5-enrollment-demographics22f9e85b8c9b66d6b292ff0000215f92.pdf?sfvrsn=2>
- Taylor, G., Shepard, L., Kinner, F., & Rosenthal, J. (2001). A survey of teachers' perspectives on high-stakes testing in Colorado: What gets taught, what gets lost. Retrieved from <http://nepc.colorado.edu/publication/a-survey-teachers-perspectives-high-stakes-testing-colorado-what-gets-taught-what-gets-lost>.
- “The data story” (2017). *New Schools for New Orleans*. Retrieved from <http://www.newschoolsforneworleans.org/education-in-nola/data-and-analysis-old/>
- 2013 mathematics and reading. (2017). *The Nation’s Report Card*. Retrieved from https://www.nationsreportcard.gov/reading_math_2013/#/achievement-gaps
- 2015 mathematics and reading assessments. (2017). *The Nation’s Report Card*. Retrieved from https://www.nationsreportcard.gov/reading_math_2015/#reading/groups?grade=4
- 2015-2017 LEAP assessment results---percent of students scored mastery and above for subgroups combining ELA, math, and science. Retrieved from <http://www.louisianabelieves.com/resources/library/test-results>
- 2017-2018 Louisiana assessment calendar. (2017). Retrieved from <http://www.louisianabelieves.com/docs/default-source/assessment/2017-2018-assessment-calendar.pdf>
- Valli, L. & Buese, D. (2007). The changing roles of teachers in an era of high-stakes accountability. *American Education Research Journal*, 44(3), 515-558.
- Vanicore, A. (2016, December). Birth of New Orleans' all-charter school system: The challenges, risks, Expectations. *The New Orleans Advocate*. Retrieved from

http://www.theadvocate.com/new_orleans/news/education/article_fe3b66fc-ca00-11e6-86f2-77f5e1d88e72.html

Viega, M. (2013). *Loving me and my butterfly wings: A study of hip-hop songs written by adolescents in music therapy*. (Unpublished doctoral dissertation). Temple University. Philadelphia, PA.

Walker, T. (2014). NEA survey: Nearly half of teachers consider leaving profession due to standardized testing. Retrieved from <http://neatoday.org/2014/11/02/nea-survey-nearly-half-of-teachers-consider-leaving-profession-due-to-standardized-testing-2/>

Willis, J. (2007). Putting the squeeze on social studies: Managing teaching dilemmas in subject areas excluded from state testing. *Teachers College Record*. 9(8), 1980-2046.

Wolcott, H. (1994). *Transforming qualitative data: Description, analysis, and interpretation*. Thousand Oaks, CA: Sage.

Wolf, N. L. (2011). A case study comparison of charter and traditional schools in New Orleans Recovery School District: Selection criteria and service provision for students with disabilities. *Remedial and Special Education*, 32(5), 382–392.

Yeh, S. (2005). Limiting the unintended consequences of high-stakes testing. *Education Policy Analysis Archives*, 13(43).

Vita

Desi Richter was born in Brewton, Alabama. She obtained her Bachelors of Music Education degree from Otterbein University in 1996. She began pursuing her Ph.D.in Curriculum and Instruction from the University of New Orleans in 2014. She is a currently co-teacher for the Greater New Orleans Writing Project and an independent academic writing consultant.